



Nadir Jeevanjee

An Introduction to Tensors and Group Theory for Physicists

 Birkhäuser

Nadir Jeevanjee

An Introduction to Tensors and Group Theory for Physicists

Nadir Jeevanjee
Department of Physics
University of California
366 LeConte Hall MC 7300
Berkeley, CA 94720
USA
jeevanje@berkeley.edu

ISBN 978-0-8176-4714-8 e-ISBN 978-0-8176-4715-5
DOI 10.1007/978-0-8176-4715-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011935949

Mathematics Subject Classification (2010): 15Axx, 20Cxx, 22Exx, 81Rxx

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

To My Parents

Preface

This book is composed of two parts: Part I (Chaps. 1 through 3) is an introduction to tensors and their physical applications, and Part II (Chaps. 4 through 6) introduces group theory and intertwines it with the earlier material. Both parts are written at the advanced-undergraduate/beginning-graduate level, although in the course of Part II the sophistication level rises somewhat. Though the two parts differ somewhat in flavor, I have aimed in both to fill a (perceived) gap in the literature by connecting the component formalisms prevalent in physics calculations to the abstract but more conceptual formulations found in the math literature. My firm belief is that we need to see tensors and groups in coordinates to get a sense of how they work, but also need an abstract formulation to understand their essential nature and organize our thinking about them.

My original motivation for the book was to demystify tensors and provide a unified framework for understanding them in all the different contexts in which they arise in physics. The word tensor is ubiquitous in physics (stress tensor, moment-of-inertia tensor, field tensor, metric tensor, tensor product, etc.) and yet tensors are rarely defined carefully, and the definition usually has to do with transformation properties, making it difficult to get a feel for what these objects *are*. Furthermore, physics texts at the beginning graduate level usually only deal with tensors in their component form, so students wonder what the difference is between a second rank tensor and a matrix, and why new, enigmatic terminology is introduced for something they have already seen. All of this produces a lingering unease, which I believe can be alleviated by formulating tensors in a more abstract but conceptually much clearer way. This coordinate-free formulation is standard in the mathematical literature on differential geometry and in physics texts on General Relativity, but as far as I can tell is not accessible to undergraduates or beginning graduate students in physics who just want to learn what a tensor is *without* dealing with the full machinery of tensor analysis on manifolds.

The irony of this situation is that a proper understanding of tensors does not require much more mathematics than what you likely encountered as an undergraduate. In Chap. 2 I introduce this additional mathematics, which is just an extension of the linear algebra you probably saw in your lower-division coursework. This material sets the stage for tensors, and hopefully also illuminates some of the more

enigmatic objects from quantum mechanics and relativity, such as bras and kets, covariant and contravariant components of vectors, and spherical harmonics. After laying the necessary linear algebraic foundations, we give in Chap. 3 the modern (component-free) definition of tensors, all the while keeping contact with the coordinate and matrix representations of tensors and their transformation laws. Applications in classical and quantum physics follow.

In Part II of the book I introduce group theory and its physical applications, which is a beautiful subject in its own right and also a nice application of the material in Part I. There are many good books on the market for group theory and physics (see the references), so rather than be exhaustive I have just attempted to present those aspects of the subject most essential for upper-division and graduate-level physics courses. In Chap. 4 I introduce abstract groups, but quickly illustrate that concept with myriad examples from physics. After all, there would be little point in making such an abstract definition if it did not subsume many cases of interest! We then introduce Lie groups and their associated Lie algebras, making precise the nature of the symmetry ‘generators’ that are so central in quantum mechanics. Much time is also spent on the groups of rotations and Lorentz transformations, since these are so ubiquitous in physics.

In Chap. 5 I introduce representation theory, which is a mathematical formalization of what we mean by the ‘transformation properties’ of an object. This subject sews together the material from Chaps. 3 and 4, and is one of the most important applications of tensors, at least for physicists. Chapter 6 then applies and extends the results of Chap. 5 to a few specific topics: the perennially mysterious ‘spherical’ tensors, the Wigner–Eckart theorem, and Dirac bilinears. The presentation of these later topics is admittedly somewhat abstract, but I believe that the mathematically precise treatment yields insights and connections not usually found in the usual physicist’s treatment of the subjects.

This text aims (perhaps naively!) to be simultaneously intuitive and rigorous. Thus, although much of the language (especially in the examples) is informal, almost all the definitions given are precise and are the same as one would find in a pure math text. This may put you off if you feel less mathematically inclined; I hope, however, that you will work through your discomfort and develop the necessary mathematical sophistication, as the results will be well worth it. Furthermore, if you can work your way through the text (or at least most of Chap. 5), you will be well prepared to tackle graduate math texts in related areas.

As for prerequisites, it is assumed that you have been through the usual undergraduate physics curriculum, including a “mathematical methods for physicists” course (with at least a cursory treatment of vectors and matrices), as well as the standard upper-division courses in classical mechanics, quantum mechanics, and relativity. Any undergraduate versed in those topics, as well as any graduate student in physics, should be able to read this text. To undergraduates who are eager to learn about tensors but have not yet completed the standard curriculum, I apologize; many of the examples and practically all of the motivation for the text come from those courses, and to assume no knowledge of those topics would preclude discussion of the many applications that motivated me to write this book in the first place.

However, if you are motivated and willing to consult the references, you could certainly work through this text, and would no doubt be in excellent shape for those upper-division courses once you take them.

Exercises and problems are included in the text, with exercises occurring within the chapters and problems occurring at the end of each chapter. The exercises in particular should be done as they arise, or at least carefully considered, as they often flesh out the text and provide essential practice in using the definitions. Very few of the exercises are computationally intensive, and many of them can be done in a few lines. They are designed primarily to test your conceptual understanding and help you internalize the subject. Please do not ignore them!

Besides the aforementioned prerequisites I have also indulged in the use of some very basic mathematical shorthand for brevity's sake; a guide is below. Also, be aware that for simplicity's sake I have set all physical constants such as c and \hbar equal to 1. Enjoy!

Berkeley, USA

Nadir Jeevanjee

Acknowledgements

Many thanks to those who have provided feedback on rough drafts of this text, particularly Hal Haggard, Noah Bray-Ali, Sourav K. Mandal, Mark Moriarty, Albert Shieh, Felicitas Hernandez, and Emily Rauscher. Apologies to anyone I've forgotten! Thanks also to the students of Phys 198 at Berkeley in the Spring of 2011, who volunteered to be experimental subjects in a course based on Chaps. 1–3; getting their perspective and rewriting accordingly has (I hope) made the book much more accessible and user friendly. Thanks also to professors Robert Penner and Ko Honda of the U.S.C. mathematics department; their encouragement and early mentorship was instrumental in my pursuing math and physics to the degree that I have. Finally, thanks to my family, whose support has never wavered, regardless of the many turns my path has taken. Onwards!

Contents

Part I Linear Algebra and Tensors

1	A Quick Introduction to Tensors	3
2	Vector Spaces	9
2.1	Definition and Examples	9
2.2	Span, Linear Independence, and Bases	14
2.3	Components	17
2.4	Linear Operators	21
2.5	Dual Spaces	25
2.6	Non-degenerate Hermitian Forms	27
2.7	Non-degenerate Hermitian Forms and Dual Spaces	31
2.8	Problems	33
3	Tensors	39
3.1	Definition and Examples	39
3.2	Change of Basis	43
3.3	Active and Passive Transformations	49
3.4	The Tensor Product—Definition and Properties	53
3.5	Tensor Products of V and V^*	55
3.6	Applications of the Tensor Product in Classical Physics	58
3.7	Applications of the Tensor Product in Quantum Physics	60
3.8	Symmetric Tensors	68
3.9	Antisymmetric Tensors	70
3.10	Problems	81

Part II Group Theory

4	Groups, Lie Groups, and Lie Algebras	87
4.1	Groups—Definition and Examples	88
4.2	The Groups of Classical and Quantum Physics	96
4.3	Homomorphism and Isomorphism	103

4.4	From Lie Groups to Lie Algebras	111
4.5	Lie Algebras—Definition, Properties, and Examples	115
4.6	The Lie Algebras of Classical and Quantum Physics	121
4.7	Abstract Lie Algebras	127
4.8	Homomorphism and Isomorphism Revisited	131
4.9	Problems	138
5	Basic Representation Theory	145
5.1	Representations: Definitions and Basic Examples	145
5.2	Further Examples	150
5.3	Tensor Product Representations	159
5.4	Symmetric and Antisymmetric Tensor Product Representations . .	165
5.5	Equivalence of Representations	169
5.6	Direct Sums and Irreducibility	177
5.7	More on Irreducibility	184
5.8	The Irreducible Representations of $\mathfrak{su}(2)$, $SU(2)$ and $SO(3)$	188
5.9	Real Representations and Complexifications	193
5.10	The Irreducible Representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, $SL(2, \mathbb{C})$ and $SO(3, 1)_o$	196
5.11	Irreducibility and the Representations of $O(3, 1)$ and Its Double Covers	204
5.12	Problems	208
6	The Wigner–Eckart Theorem and Other Applications	213
6.1	Tensor Operators, Spherical Tensors and Representation Operators	213
6.2	Selection Rules and the Wigner–Eckart Theorem	217
6.3	Gamma Matrices and Dirac Bilinears	222
6.4	Problems	225
Appendix	Complexifications of Real Lie Algebras and the Tensor Product Decomposition of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ Representations	227
A.1	Direct Sums and Complexifications of Lie Algebras	227
A.2	Representations of Complexified Lie Algebras and the Tensor Product Decomposition of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ Representations	229
References	235
Index	237

Notation

Some Mathematical Shorthand

\mathbb{R}	The set of real numbers
\mathbb{C}	The set of complex numbers
\mathbb{Z}	The set of positive and negative integers
\in	“is an element of”, “an element of”, i.e. $2 \in \mathbb{R}$ reads “2 is an element of the real numbers”
\notin	“is not an element of”
\forall	“for all”
\subset	“is a subset of”, “a subset of”
\equiv	Denotes a definition
$f : A \rightarrow B$	Denotes a map f that takes elements of the set A into elements of the set B
$f : a \mapsto b$	Indicates that the map f sends the element a to the element b
\circ	Denotes a composition of maps, i.e. if $f : A \rightarrow B$ and $g : B \rightarrow C$, then $g \circ f : A \rightarrow C$ is given by $(g \circ f)(a) \equiv g(f(a))$
$A \times B$	The set $\{(a, b)\}$ of all ordered pairs where $a \in A$, $b \in B$. Referred to as the <i>cartesian product</i> of sets A and B . Extends in the obvious way to n -fold products $A_1 \times \cdots \times A_n$
\mathbb{R}^n	$\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}$
\mathbb{C}^n	$\underbrace{\mathbb{C} \times \cdots \times \mathbb{C}}_{n \text{ times}}$
$\{A \mid Q\}$	Denotes a set A subject to condition Q . For instance, the set of all even integers can be written as $\{x \in \mathbb{R} \mid x/2 \in \mathbb{Z}\}$
\square	Denotes the end of a proof or example

*Dirac Dictionary*¹

Standard Notation Dirac Notation

Vector $\psi \in \mathcal{H}$	$ \psi\rangle$
Dual Vector $L(\psi)$	$\langle\psi $
Inner Product (ψ, ϕ)	$\langle\psi \phi\rangle$
$A(\psi), A \in \mathcal{L}(\mathcal{H})$	$A \psi\rangle$
$(\psi, A\phi)$	$\langle\psi A \phi\rangle$
$T_i{}^j e^i \otimes e_j$	$\sum_{i,j} T_{ij} j\rangle\langle i $
$e_i \otimes e_j$	$ i\rangle j\rangle$ or $ i, j\rangle$

¹We summarize here all of the translations given in the text between quantum-mechanical Dirac notation and standard mathematical notation.

Part I
Linear Algebra and Tensors

Chapter 1

A Quick Introduction to Tensors

The reason tensors are introduced in a somewhat ad-hoc manner in most physics courses is twofold: first, a detailed and proper understanding of tensors requires mathematics that is slightly more abstract than the standard linear algebra and vector calculus that physics students use everyday. Second, students do not necessarily *need* such an understanding to be able to manipulate tensors and solve problems with them. The drawback, of course, is that many students feel uneasy whenever tensors are discussed, and they find that they can use tensors for computation but do not have an intuitive feel for what they are doing. One of the primary aims of this book is to alleviate those feelings. Doing that, however, requires a modest investment (about 30 pages) in some abstract linear algebra, so before diving into the details we will begin with a rough overview of what a tensor is, which hopefully will whet your appetite and tide you over until we can discuss tensors in full detail in Chap. 3.

Many older books define a tensor as a collection of objects which carry indices and which ‘transform’ in a particular way specified by those indices. Unfortunately, this definition usually does not yield much insight into what a tensor is. One of the main purposes of the present text is to promulgate the more modern definition of a tensor, which is equivalent to the old one but is more conceptual and is in fact already standard in the mathematics literature. This definition takes a tensor to be a *function* which eats a certain number of vectors (known as the **rank** r of the tensor) and produces a number. The distinguishing characteristic of a tensor is a special property called **multilinearity**, which means that it must be linear in each of its r arguments (recall that linearity for a function with a single argument just means that $T(v + cw) = T(v) + cT(w)$ for all vectors v and w and numbers c). As we will explain in a moment, this multilinearity enables us to express the value of the function on an *arbitrary* set of r vectors in terms of the values of the function on r *basis* vectors like \hat{x} , \hat{y} , and \hat{z} . These values of the function on basis vectors are nothing but the familiar **components** of the tensor, which in older treatments are usually introduced first as part of the definition of the tensor.

To make this concrete, consider a rank 2 tensor T , whose job it is to eat two vectors v and w and produce a number which we will denote as $T(v, w)$. For such a tensor, multilinearity means

$$T(v_1 + cv_2, w) = T(v_1, w) + cT(v_2, w) \quad (1.1)$$

$$T(v, w_1 + cw_2) = T(v, w_1) + cT(v, w_2) \quad (1.2)$$

for any number c and all vectors v and w . This means that if we have a coordinate basis for our vector space, say $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$, then T is determined entirely by its values on the basis vectors, as follows: first, expand v and w in the coordinate basis as

$$\begin{aligned} v &= v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} + v_z \hat{\mathbf{z}} \\ w &= w_x \hat{\mathbf{x}} + w_y \hat{\mathbf{y}} + w_z \hat{\mathbf{z}}. \end{aligned}$$

Then by (1.1) and (1.2) we have

$$\begin{aligned} T(v, w) &= T(v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} + v_z \hat{\mathbf{z}}, w_x \hat{\mathbf{x}} + w_y \hat{\mathbf{y}} + w_z \hat{\mathbf{z}}) \\ &= v_x T(\hat{\mathbf{x}}, w_x \hat{\mathbf{x}} + w_y \hat{\mathbf{y}} + w_z \hat{\mathbf{z}}) + v_y T(\hat{\mathbf{y}}, w_x \hat{\mathbf{x}} + w_y \hat{\mathbf{y}} + w_z \hat{\mathbf{z}}) \\ &\quad + v_z T(\hat{\mathbf{z}}, w_x \hat{\mathbf{x}} + w_y \hat{\mathbf{y}} + w_z \hat{\mathbf{z}}) \\ &= v_x w_x T(\hat{\mathbf{x}}, \hat{\mathbf{x}}) + v_x w_y T(\hat{\mathbf{x}}, \hat{\mathbf{y}}) + v_x w_z T(\hat{\mathbf{x}}, \hat{\mathbf{z}}) + v_y w_x T(\hat{\mathbf{y}}, \hat{\mathbf{x}}) \\ &\quad + v_y w_y T(\hat{\mathbf{y}}, \hat{\mathbf{y}}) + v_y w_z T(\hat{\mathbf{y}}, \hat{\mathbf{z}}) + v_z w_x T(\hat{\mathbf{z}}, \hat{\mathbf{x}}) + v_z w_y T(\hat{\mathbf{z}}, \hat{\mathbf{y}}) \\ &\quad + v_z w_z T(\hat{\mathbf{z}}, \hat{\mathbf{z}}). \end{aligned}$$

If we then abbreviate our notation as

$$\begin{aligned} T_{xx} &\equiv T(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \\ T_{xy} &\equiv T(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \\ T_{yx} &\equiv T(\hat{\mathbf{y}}, \hat{\mathbf{x}}) \end{aligned} \quad (1.3)$$

and so on, we have

$$\begin{aligned} T(v, w) &= v_x w_x T_{xx} + v_x w_y T_{xy} + v_x w_z T_{xz} + v_y w_x T_{yx} + v_y w_y T_{yy} \\ &\quad + v_y w_z T_{yz} + v_z w_x T_{zx} + v_z w_y T_{zy} + v_z w_z T_{zz} \end{aligned} \quad (1.4)$$

which may look familiar from discussion of tensors in the physics literature. In that literature, the above equation is often part of the *definition* of a second rank tensor; here, though, we see that its form is really just a consequence of multilinearity. Another advantage of our approach is that the **components** $\{T_{xx}, T_{xy}, T_{xz}, \dots\}$ of T have a meaning beyond that of just being coefficients that appear in expressions like (1.4); from (1.3), we see that **components are the values of the tensor when evaluated on a given set of basis vectors**. This fact is crucial in getting a feel for tensors and what they mean.

Another nice feature of our definition of a tensor is that it allows us to *derive* the tensor transformation laws which historically were taken as the definition of a tensor. Say we switch to a new set of basis vectors $\{\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}'\}$ which are related to the old basis vectors by

$$\begin{aligned} \hat{\mathbf{x}}' &= A_{x'x} \hat{\mathbf{x}} + A_{x'y} \hat{\mathbf{y}} + A_{x'z} \hat{\mathbf{z}} \\ \hat{\mathbf{y}}' &= A_{y'x} \hat{\mathbf{x}} + A_{y'y} \hat{\mathbf{y}} + A_{y'z} \hat{\mathbf{z}} \\ \hat{\mathbf{z}}' &= A_{z'x} \hat{\mathbf{x}} + A_{z'y} \hat{\mathbf{y}} + A_{z'z} \hat{\mathbf{z}}. \end{aligned} \quad (1.5)$$

This does not affect the action of T , since T **exists independently of any basis**, but if we would like to compute the value of $T(v, w)$ in terms of the new components (v'_x, v'_y, v'_z) and (w'_x, w'_y, w'_z) of v and w , then we will need to know what the new components $\{T_{x'x'}, T_{x'y'}, T_{x'z'}, \dots\}$ look like. Computing $T_{x'x'}$, for instance, gives

$$\begin{aligned}
 T_{x'x'} &= T(\hat{\mathbf{x}}', \hat{\mathbf{x}}') \\
 &= T(A_{x'x}\hat{\mathbf{x}} + A_{x'y}\hat{\mathbf{y}} + A_{x'z}\hat{\mathbf{z}}, A_{x'x}\hat{\mathbf{x}} + A_{x'y}\hat{\mathbf{y}} + A_{x'z}\hat{\mathbf{z}}) \\
 &= A_{x'x}A_{x'x}T_{xx} + A_{x'y}A_{x'x}T_{yx} + A_{x'z}A_{x'x}T_{zx} + A_{x'x}A_{x'y}T_{xy} \\
 &\quad + A_{x'y}A_{x'y}T_{yy} + A_{x'z}A_{x'y}T_{zy} + A_{x'x}A_{x'z}T_{xz} + A_{x'y}A_{x'z}T_{yz} \\
 &\quad + A_{x'z}A_{x'z}T_{zz}.
 \end{aligned} \tag{1.6}$$

You probably recognize this as an instance of the standard tensor transformation law, which used to be taken as the definition of a tensor. Here, the transformation law is another consequence of multilinearity! In Chap. 3 we will introduce more convenient notation, which will allow us to use the Einstein summation convention, so that we can write the general form of (1.6) as

$$T_{i'j'} = A_{i'}^k A_{j'}^l T_{kl}, \tag{1.7}$$

a form which may be more familiar to some readers.

One common source of confusion is that in physics textbooks, tensors (usually of the second rank) are often represented as matrices, so then the student wonders what the difference is between a matrix and a tensor. Above, we have defined a tensor as a multilinear function on a vector space. What does that have to do with a matrix? Well, *if* we choose a basis $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$, we can then write the corresponding components $\{T_{xx}, T_{xy}, T_{xz}, \dots\}$ in the form of a matrix $[T]$ as

$$[T] \equiv \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix}. \tag{1.8}$$

Equation (1.4) can then be written compactly as

$$T(v, w) = (v_x, v_y, v_z) \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \tag{1.9}$$

where the usual matrix multiplication is implied. Thus, once a basis is chosen, the action of T can be neatly expressed using the corresponding matrix $[T]$. It is crucial to keep in mind, though, that this association between a tensor and a matrix depends entirely on a *choice of basis*, and that $[T]$ is useful mainly as a computational tool, not a conceptual handle. T is best thought of abstractly as a multilinear function, and $[T]$ as its representation in a particular coordinate system.

One possible objection to our approach is that matrices and tensors are often thought of as linear operators which take vectors into vectors, as opposed to objects which eat vectors and spit out numbers. It turns out, though, that for a second rank

tensor these two notions are equivalent. Say we have a linear operator R ; then we can turn R into a second rank tensor T_R by

$$T_R(v, w) \equiv v \cdot R w \quad (1.10)$$

where \cdot denotes the usual dot product of vectors. You can roughly interpret this tensor as taking in v and w and giving back the “component of $R w$ lying along v ”. You can also easily check that T_R is multilinear, i.e. it satisfies (1.1) and (1.2). If we compute the components of T_R we find that, for instance,

$$\begin{aligned} (T_R)_{xx} &= T_R(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \\ &= \hat{\mathbf{x}} \cdot R \hat{\mathbf{x}} \\ &= \hat{\mathbf{x}} \cdot (R_{xx} \hat{\mathbf{x}} + R_{xy} \hat{\mathbf{y}} + R_{xz} \hat{\mathbf{z}}) \\ &= R_{xx} \end{aligned}$$

so the components of the tensor T_R are the same as the components of the linear operator R ! In components, the action of T_R then looks like

$$T_R(v, w) = (v_x, v_y, v_z) \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} \quad (1.11)$$

which is identical to (1.9). This makes it obvious how to turn a linear operator R into a second rank tensor—just sandwich the component matrix in between two vectors! This whole process is also reversible: we can turn a second rank tensor T into a linear operator R_T by defining the components of the vector $R_T(v)$ as

$$(R_T(v))_x \equiv T(v, \hat{\mathbf{x}}), \quad (1.12)$$

and similarly for the other components. One can check that these processes are inverses of each other, so this sets up a one-to-one correspondence between linear operators and second rank tensors and we can thus regard them as equivalent. Since the matrix form of both is identical, one often does not bother trying to clarify exactly which one is at hand, and often times it is just a matter of interpretation.

How does all this work in a physical context? One nice example is the rank 2 moment-of-inertia tensor \mathcal{I} , familiar from rigid body dynamics in classical mechanics. In most textbooks, this tensor usually arises when one considers the relationship between the angular velocity vector ω and the angular momentum L or kinetic energy KE of a rigid body. If one chooses a basis $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ and then expresses, say, KE in terms of ω , one gets an expression like (1.4) with $v = w = \omega$ and $T_{ij} = \mathcal{I}_{ij}$. From this expression, most textbooks then figure out how the \mathcal{I}_{ij} must transform under a change of coordinates, and this behavior under coordinate change is then what identifies the \mathcal{I}_{ij} as a tensor. In this text, though, we will take a different point of view. We will *define* \mathcal{I} to be the function which, for a given rigid body, assigns to a state of rotation (described by the angular momentum vector ω) twice the corresponding kinetic energy KE . One can then calculate KE in terms of ω , which yields an expression of the form (1.4); this then shows that \mathcal{I} is a tensor in the old sense of the

term, which also means that it is a tensor in the modern sense of being a multilinear function. We can then think of \mathcal{I} as a second rank tensor defined by

$$KE = \frac{1}{2} \mathcal{I}(\omega, \omega). \quad (1.13)$$

From our point of view, we see that the components of \mathcal{I} have a physical meaning; for instance, $\mathcal{I}_{xx} = \mathcal{I}(\hat{\mathbf{x}}, \hat{\mathbf{x}})$ is just twice the kinetic energy of the rigid body when $\omega = \hat{\mathbf{x}}$. When one actually needs to compute the kinetic energy for a rigid body, one usually picks a convenient coordinate system, computes the components of \mathcal{I} using the standard formulas,¹ and then uses the convenient matrix representation to compute

$$KE = \frac{1}{2} (\omega_x, \omega_y, \omega_z) \begin{pmatrix} \mathcal{I}_{xx} & \mathcal{I}_{xy} & \mathcal{I}_{xz} \\ \mathcal{I}_{yx} & \mathcal{I}_{yy} & \mathcal{I}_{yz} \\ \mathcal{I}_{zx} & \mathcal{I}_{zy} & \mathcal{I}_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}, \quad (1.14)$$

which is the familiar expression often found in mechanics texts.

As mentioned above, one can also interpret \mathcal{I} as a linear operator which takes vectors into vectors. If we let \mathcal{I} act on the angular velocity vector, we get

$$\begin{pmatrix} \mathcal{I}_{xx} & \mathcal{I}_{xy} & \mathcal{I}_{xz} \\ \mathcal{I}_{yx} & \mathcal{I}_{yy} & \mathcal{I}_{yz} \\ \mathcal{I}_{zx} & \mathcal{I}_{zy} & \mathcal{I}_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \quad (1.15)$$

which you probably recognize as the coordinate expression for the angular momentum L . Thus, \mathcal{I} can be interpreted *either* as the rank 2 tensor which eats two copies of the angular velocity vector ω and produces the kinetic energy, or as the linear operator which takes ω into L . The two definitions are equivalent.

Many other tensors which we will consider in this text are also nicely understood as multilinear functions which happen to have convenient component and matrix representations. These include the electromagnetic field tensor of electrodynamics, as well as the metric tensors of Newtonian and relativistic mechanics. As we progress we will see how we can use our new point of view to get a handle on these usually somewhat enigmatic objects.

¹See Example 3.14 or any standard textbook such as Goldstein [6].

Chapter 2

Vector Spaces

Since tensors are a special class of functions defined on vector spaces, we must have a good foundation in linear algebra before discussing them. In particular, you will need a little bit more linear algebra than is covered in most sophomore or junior level linear algebra/ODE courses. This chapter starts with the familiar material about vectors, bases, linear operators etc. but eventually moves on to slightly more sophisticated topics that are essential for understanding tensors in physics. As we lay this groundwork, hopefully you will find that our slightly more abstract viewpoint also clarifies the nature of many of the objects in physics you have already encountered.

2.1 Definition and Examples

We begin with the definition of an abstract vector space. We are taught as undergraduates to think of vectors as arrows with a head and a tail, or as ordered triples of real numbers, but physics, and especially quantum mechanics, requires a more abstract notion of vectors. Before reading the definition of an abstract vector space, keep in mind that the definition is supposed to distill all the essential features of vectors as we know them (like addition and scalar multiplication) while detaching the notion of a vector space from specific constructs, like ordered n -tuples of real or complex numbers (denoted as \mathbb{R}^n and \mathbb{C}^n , respectively). The mathematical utility of this is that much of what we know about vector spaces depends only on the essential properties of addition and scalar multiplication, not on other properties particular to \mathbb{R}^n or \mathbb{C}^n . If we work in the abstract framework and then come across other mathematical objects that do not look like \mathbb{R}^n or \mathbb{C}^n but that *are* abstract vector spaces, then most everything we know about \mathbb{R}^n and \mathbb{C}^n will apply to these spaces as well. Physics also forces us to use the abstract definition since many quantum-mechanical vector spaces are infinite-dimensional and cannot be viewed as \mathbb{C}^n or \mathbb{R}^n for any n . An added dividend of the abstract approach is that we will learn to think about vector spaces independently of any basis, which will prove very useful.

That said, an (*abstract*) **vector space** is a set V (whose elements are called vectors), together with a set of scalars C (for us, C is always \mathbb{R} or \mathbb{C}) and operations of addition and scalar multiplication that satisfy the following axioms:

1. $v + w = w + v$ for all v, w in V (Commutativity)
2. $v + (w + x) = (v + w) + x$ for all v, w, x in V (Associativity)
3. There exists a vector 0 in V such that $v + 0 = v$ for all v in V
4. For all v in V there is a vector $-v$ such that $v + (-v) = 0$
5. $c(v + w) = cv + cw$ for all v and w in V and scalars c (Distributivity)
6. $1v = v$ for all v in V
7. $(c_1 + c_2)v = c_1v + c_2v$ for all scalars c_1, c_2 and vectors v
8. $(c_1c_2)v = c_1(c_2v)$ for all scalars c_1, c_2 and vectors v

Some parts of the definition may seem tedious or trivial, but they are just meant to ensure that the addition and scalar multiplication operations behave the way we expect them to. In determining whether a set is a vector space or not, one is usually most concerned with defining addition in such a way that the set is closed under addition and that axioms 3 and 4 are satisfied; most of the other axioms are so natural and obviously satisfied that one, in practice, rarely bothers to check them.¹ That said, let us look at some examples from physics, most of which will recur throughout the text.

Example 2.1 \mathbb{R}^n

This is the most basic example of a vector space, and the one on which the abstract definition is modeled. Addition and scalar multiplication are defined in the usual way: for $v = (v^1, v^2, \dots, v^n)$, $w = (w^1, w^2, \dots, w^n)$ in \mathbb{R}^n , we have

$$(v^1, v^2, \dots, v^n) + (w^1, w^2, \dots, w^n) = (v^1 + w^1, v^2 + w^2, \dots, v^n + w^n)$$

and

$$c(v^1, v^2, \dots, v^n) = (cv^1, cv^2, \dots, cv^n).$$

You should check that the axioms are satisfied. These spaces, of course, are basic in physics; \mathbb{R}^3 is our usual three-dimensional cartesian space, \mathbb{R}^4 is spacetime in special relativity, and \mathbb{R}^n for higher n occurs in classical physics as configuration spaces for multiparticle systems (for example, \mathbb{R}^6 is the configuration space in the

¹Another word about axioms 3 and 4, for the mathematically inclined (feel free to skip this if you like): the axioms do not demand that the zero element and inverses are unique, but this actually follows easily *from* the axioms. If 0 and $0'$ are two zero elements, then

$$0 = 0 + 0' = 0'$$

and so the zero element is unique. Similarly, if $-v$ and $-v'$ are both inverse to some vector v , then

$$-v' = -v' + 0 = -v' + (v - v) = (-v' + v) - v = 0 - v = -v$$

and so inverses are unique as well.

classic two-body problem, as you need six coordinates to specify the position of two particles in three-dimensional space).

Example 2.2 \mathbb{C}^n

This is another basic example—addition and scalar multiplication are defined as for \mathbb{R}^n , and the axioms are again straightforward to verify. Note, however, that \mathbb{C}^n is a **complex** vector space, i.e. the set C in the definition is \mathbb{C} so scalar multiplication by complex numbers is defined, whereas \mathbb{R}^n is only a **real** vector space. This seemingly pedantic distinction can often end up being significant, as we will see.² \mathbb{C}^n occurs in physics primarily as the ket space for finite-dimensional quantum-mechanical systems, such as particles with spin but without translational degrees of freedom. For instance, a spin 1/2 particle fixed in space has ket space identifiable with \mathbb{C}^2 , and more generally a fixed particle with spin s has ket space identifiable with \mathbb{C}^{2s+1} .

Example 2.3 $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$, $n \times n$ matrices with real or complex entries

The vector space structure of $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ is similar to that of \mathbb{R}^n and \mathbb{C}^n : denoting the entry in the i th row and j th column of a matrix A as A_{ij} , we define addition and (real) scalar multiplication for $A, B \in M_n(\mathbb{R})$ by

$$\begin{aligned}(A + B)_{ij} &= A_{ij} + B_{ij} \\ (cA)_{ij} &= cA_{ij}\end{aligned}$$

i.e. addition and scalar multiplication are done component-wise. The same definitions are used for $M_n(\mathbb{C})$, which is of course a complex vector space. You can again check that the axioms are satisfied. Though these vector spaces do not appear explicitly in physics very often, they have many important subspaces, one of which we consider in the next example.

Example 2.4 $H_n(\mathbb{C})$, $n \times n$ Hermitian matrices with complex entries

$H_n(\mathbb{C})$, the set of all $n \times n$ Hermitian matrices,³ is obviously a subset of $M_n(\mathbb{C})$, and in fact it is a *subspace* of $M_n(\mathbb{C})$ in that it forms a vector space itself.⁴ To show this it is not necessary to verify all of the axioms, since most of them are satisfied by virtue of $H_n(\mathbb{C})$ being a subset of $M_n(\mathbb{C})$; for instance, addition and scalar multiplication in $H_n(\mathbb{C})$ are just given by the restriction of those operations in $M_n(\mathbb{C})$ to $H_n(\mathbb{C})$, so

²Note also that any complex vector space is also a real vector space, since if you know how to multiply vectors by a complex number, then you certainly know how to multiply them by a real number. The converse, of course, is not true.

³Hermitian matrices being those which satisfy $A^\dagger \equiv (A^T)^* = A$, where superscript T denotes the transpose and superscript $*$ denotes complex conjugation of the entries.

⁴Another footnote for the mathematically inclined: as discussed later in this example, though, $H_n(\mathbb{C})$ is only a *real* vector space, so it is only a subspace of $M_n(\mathbb{C})$ when $M_n(\mathbb{C})$ is considered as a real vector space, cf. footnote 2.

the commutativity of addition and the distributivity of scalar multiplication over addition follow immediately. What does remain to be checked is that $H_n(\mathbb{C})$ is closed under addition and contains the zero “vector” (in this case, the zero matrix), both of which are easily verified. One interesting thing about $H_n(\mathbb{C})$ is that even though the entries of its matrices can be complex, it does *not* form a complex vector space; multiplying a Hermitian matrix by i yields an *anti*-Hermitian matrix, so $H_n(\mathbb{C})$ is *not* closed under complex scalar multiplication. As far as physical applications go, we know that physical observables in quantum mechanics are represented by Hermitian operators, and if we are dealing with a finite-dimensional ket space such as those mentioned in Example 2.2 then observables can be represented as elements of $H_n(\mathbb{C})$. As an example one can take a fixed spin $1/2$ particle whose ket space is \mathbb{C}^2 ; the angular momentum operators are then represented as $L_i = \frac{1}{2}\sigma_i$, where the σ_i are the Hermitian **Pauli matrices**

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.1)$$

Example 2.5 $L^2([a, b])$, square-integrable complex-valued functions on an interval

This example is fundamental in quantum mechanics. A complex-valued function f on $[a, b] \subset \mathbb{R}$ is said to be **square-integrable** if

$$\int_a^b |f(x)|^2 dx < \infty. \quad (2.2)$$

Defining addition and scalar multiplication in the obvious way,

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ (cf)(x) &= cf(x), \end{aligned}$$

and taking the zero element to be the function which is identically zero (i.e. $f(x) = 0$ for all x) yields a complex vector space. (Note that if we considered only real-valued functions then we would only have a real vector space.) Verifying the axioms is straightforward though not entirely trivial, as one must show that the sum of two square-integrable functions is again square-integrable (Problem 2.1). This vector space arises in quantum mechanics as the set of *normalizable* wavefunctions for a particle in a one-dimensional infinite potential well. Later on we will consider the more general scenario where the particle may be unbound, in which case $a = -\infty$ and $b = \infty$ and the above definitions are otherwise unchanged. This vector space is denoted as $L^2(\mathbb{R})$.

Example 2.6 $\mathcal{H}_l(\mathbb{R}^3)$ and $\tilde{\mathcal{H}}_l$, the harmonic polynomials and the spherical harmonics

Consider the set $P_l(\mathbb{R}^3)$ of all complex-coefficient polynomial functions on \mathbb{R}^3 of fixed degree l , i.e. all linear combinations of functions of the form $x^i y^j z^k$ where $i + j + k = l$. Addition and (complex) scalar multiplication are defined in the usual

way and the axioms are again easily verified, so $P_l(\mathbb{R}^3)$ is a vector space. Now consider the vector subspace $\mathcal{H}_l(\mathbb{R}^3) \subset P_l(\mathbb{R}^3)$ of **harmonic** degree l polynomials, i.e. degree l polynomials satisfying $\Delta f = 0$, where Δ is the usual three-dimensional Laplacian. You may be surprised to learn that the spherical harmonics of degree l are essentially elements of $\mathcal{H}_l(\mathbb{R}^3)$! To see the connection, note that if we write a degree l polynomial

$$f(x, y, z) = \sum_{\substack{i,j,k \\ i+j+k=l}} c_{ijk} x^i y^j z^k$$

in spherical coordinates with polar angle θ and azimuthal angle ϕ , we will get

$$f(r, \theta, \phi) = r^l Y(\theta, \phi)$$

for some function $Y(\theta, \phi)$, which is just the restriction of f to the unit sphere. If we write the Laplacian out in spherical coordinates we get

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{S^2} \quad (2.3)$$

where Δ_{S^2} is shorthand for the angular part of the Laplacian.⁵ You will show below that applying this to $f(r, \theta, \phi) = r^l Y(\theta, \phi)$ and demanding that f be harmonic yields

$$\Delta_{S^2} Y(\theta, \phi) = -l(l+1)Y(\theta, \phi), \quad (2.5)$$

which is the definition of a spherical harmonic of degree l ! Conversely, if we take a degree l spherical harmonic and multiply it by r^l , the result is a harmonic function. If we let $\tilde{\mathcal{H}}_l$ denote the set of all spherical harmonics of degree l , then we have established a one-to-one correspondence between $\tilde{\mathcal{H}}_l$ and $\mathcal{H}_l(\mathbb{R}^3)$! The correspondence is given by restricting functions in $\mathcal{H}_l(\mathbb{R}^3)$ to the unit sphere, and conversely by multiplying functions in $\tilde{\mathcal{H}}_l$ by r^l , i.e.

$$\begin{aligned} \mathcal{H}_l(\mathbb{R}^3) &\xleftrightarrow{1-1} \tilde{\mathcal{H}}_l \\ f &\longrightarrow f(r=1, \theta, \phi) \\ r^l Y(\theta, \phi) &\longleftarrow Y(\theta, \phi). \end{aligned}$$

In particular, this means that **the familiar spherical harmonics $Y_m^l(\theta, \phi)$ are just the restriction of particular harmonic degree l polynomials to the unit**

⁵The differential operator Δ_{S^2} is also sometimes known as the **spherical Laplacian**, and is given explicitly by

$$\Delta_{S^2} = \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}. \quad (2.4)$$

We will not need the explicit form of Δ_{S^2} here. A derivation and further discussion can be found in any electrodynamics or quantum mechanics book, like Sakurai [14].

sphere. For instance, consider the case $l = 1$. Clearly $\mathcal{H}_1(\mathbb{R}^3) = P_1(\mathbb{R}^3)$ since all first-degree (linear) functions are harmonic. If we write the functions

$$\frac{x + iy}{\sqrt{2}}, z, \frac{x - iy}{\sqrt{2}} \in \mathcal{H}_1(\mathbb{R}^3)$$

in spherical coordinates, we get

$$\frac{1}{\sqrt{2}}re^{i\phi}\sin\theta, r\cos\theta, \frac{1}{\sqrt{2}}re^{-i\phi}\sin\theta \in \mathcal{H}_1(\mathbb{R}^3)$$

which when restricted to the unit sphere yield

$$\frac{1}{\sqrt{2}}e^{i\phi}\sin\theta, \cos\theta, \frac{1}{\sqrt{2}}e^{-i\phi}\sin\theta \in \tilde{\mathcal{H}}_1.$$

Up to (overall) normalization, these are the usual degree 1 spherical harmonics Y_m^1 , $-1 \leq m \leq 1$. The $l = 2$ case is treated in Exercise 2.5 below. Spherical harmonics are discussed further throughout this text; for a complete discussion, see Sternberg [16].

Exercise 2.1 Verify (2.5).

Exercise 2.2 Consider the functions

$$\frac{1}{2}(x + iy)^2, z(x + iy), \frac{1}{\sqrt{2}}(x^2 + y^2 - 2z^2), z(x - iy), \frac{1}{2}(x - iy)^2 \in P_2(\mathbb{R}^3). \quad (2.6)$$

Verify that they are in fact harmonic, and then write them in spherical coordinates and restrict to the unit sphere to obtain, up to normalization, the familiar degree 2 spherical harmonics Y_m^2 , $-2 \leq m \leq 2$.

Non-example $GL(n, \mathbb{R})$, invertible $n \times n$ matrices

The ‘general linear group’ $GL(n, \mathbb{R})$, defined to be the subset of $M_n(\mathbb{R})$ consisting of invertible $n \times n$ matrices, is not a vector space though it seems like it could be. Why not?

2.2 Span, Linear Independence, and Bases

The notion of a basis is probably familiar to most readers, at least intuitively: it is a set of vectors out of which we can ‘make’ all the other vectors in a given vector space V . In this section we will make this idea precise and describe bases for some of the examples in the previous section.

First, we need the notion of the span of a set of vectors. If $S = \{v_1, v_2, \dots, v_k\} \subset V$ is a set of k vectors in V , then the **span** of S , denoted $\text{Span}\{v_1, v_2, \dots, v_k\}$ or $\text{Span } S$, is defined to be just the set of all vectors of the form $c^1v_1 + c^2v_2 + \dots + c^kv_k$. Such vectors are known as **linear combinations** of the v_i , so $\text{Span } S$ is just the set of all linear combinations of the vectors in S . For instance, if $S = \{(1, 0, 0), (0, 1, 0)\} \subset$

\mathbb{R}^3 , then $\text{Span } S$ is just the set of all vectors of the form $(c^1, c^2, 0)$ with $c^1, c^2 \in \mathbb{R}$. If S has infinitely many elements then the span of S is again all the linear combinations of vectors in S , though in this case the linear combinations can have an arbitrarily large (but finite) number of terms.⁶

Next we need the notion of linear dependence: a (not necessarily finite) set of vectors S is said to be **linearly dependent** if there exists distinct vectors v_1, v_2, \dots, v_m in S and scalars c^1, c^2, \dots, c^m , not all of which are 0, such that

$$c^1 v_1 + c^2 v_2 + \dots + c^m v_m = 0. \quad (2.7)$$

What this definition really means is that at least one vector in S can be written as a linear combination of the others, and in that sense is dependent (you should take a second to convince yourself of this). If S is not linearly dependent then we say it is **linearly independent**, and in this case no vector in S can be written as a linear combination of any others. For instance, the set $S = \{(1, 0, 0), (0, 1, 0), (1, 1, 0)\} \subset \mathbb{R}^3$ is linearly dependent whereas the set $S' = \{(1, 0, 0), (0, 1, 0), (0, 1, 1)\}$ is linearly independent, as you can check.

With these definitions in place we can now define a **basis** for a vector space V as an *ordered* linearly independent set $\mathcal{B} \subset V$ whose span is all of V . This means, roughly speaking, that a basis has enough vectors to make all the others, but no more than that. When we say that $\mathcal{B} = \{v_1, \dots, v_k\}$ is an *ordered* set we mean that the order of the v_i is part of the definition of \mathcal{B} , so another basis with the same vectors but a different order is considered inequivalent. The reasons for this will become clear as we progress.

One can show⁷ that all finite bases must have the same number of elements, so we define the **dimension** of a vector space V , denoted $\dim V$, to be the number of elements of any finite basis. If no finite basis exists, then we say that V is **infinite-dimensional**.

Also, we should mention that basis vectors are often denoted e_i rather than v_i , and we will use this notation from now on.

Exercise 2.3 Given a vector v and a finite basis $\mathcal{B} = \{e_i\}_{i=1, \dots, n}$, show that the expression of v as a linear combination of the e_i is unique.

Example 2.7 \mathbb{R}^n and \mathbb{C}^n

\mathbb{R}^n has the following natural basis, also known as the **standard basis**:

$$\{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, \dots, 1)\}.$$

You should check that this is indeed a basis, and thus that the dimension of \mathbb{R}^n is, unsurprisingly, n . The same set serves as a basis for \mathbb{C}^n , where of course

⁶We do not generally consider infinite linear combinations like $\sum_{i=1}^{\infty} c^i v_i = \lim_{N \rightarrow \infty} \sum_{i=1}^N c^i v_i$ because in that case we would need to consider whether the limit exists, i.e. whether the sum converges in some sense. More on this later.

⁷See Hoffman and Kunze [10].

now the linear combination coefficients c^i are allowed to be complex numbers. Note that although this basis is the most natural, there are infinitely many other perfectly respectable bases out there; you should check, for instance, that $\{(1, 1, 0, \dots, 0), (0, 1, 1, 0, \dots, 0), \dots, (0, \dots, 1, 1), (1, 0, \dots, 0, 1)\}$ is also a basis when $n > 2$.

Example 2.8 $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$

Let E_{ij} be the $n \times n$ matrix with a 1 in the i th row, j th column and zeros everywhere else. Then you can check that $\{E_{ij}\}_{i,j=1,\dots,n}$ is a basis for both $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$, and that both spaces have dimension n^2 . Again, there are other nice bases out there; for instance, the symmetric matrices $S_{ij} \equiv E_{ij} + E_{ji}$, $i \leq j$, and antisymmetric matrices $A_{ij} \equiv E_{ij} - E_{ji}$, $i < j$ taken together also form a basis for both $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$.

Exercise 2.4 Let $S_n(\mathbb{R})$, $A_n(\mathbb{R})$ be the sets of $n \times n$ symmetric and antisymmetric matrices, respectively. Show that both are real vector spaces, compute their dimensions, and check that $\dim S_n(\mathbb{R}) + \dim A_n(\mathbb{R}) = \dim M_n(\mathbb{R})$, as expected.

Example 2.9 $H_2(\mathbb{C})$

Let us find a basis for $H_2(\mathbb{C})$. First, we need to know what a general element of $H_2(\mathbb{C})$ looks like. In terms of complex components, the condition $A = A^\dagger$ reads

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} \quad (2.8)$$

where the bar denotes complex conjugation. This means that $a, d \in \mathbb{R}$ and $b = \bar{c}$, so in terms of real numbers we can write a general element of $H_2(\mathbb{C})$ as

$$\begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} = tI + x\sigma_x + y\sigma_y + z\sigma_z \quad (2.9)$$

where I is the identity matrix and $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices defined in (2.1). You can easily check that the set $\mathcal{B} = \{I, \sigma_x, \sigma_y, \sigma_z\}$ is linearly independent, and since (2.9) shows that \mathcal{B} spans $H_2(\mathbb{C})$, \mathcal{B} is a basis for $H_2(\mathbb{C})$. We also see that $H_2(\mathbb{C})$ has (real) dimension 4.

Exercise 2.5 Using the matrices S_{ij} and A_{ij} from Example 2.8, construct a basis for $H_n(\mathbb{C})$ and compute its dimension.

Example 2.10 $Y_m^l(\theta, \phi)$

We saw in the previous section that the Y_m^l are elements of $\tilde{\mathcal{H}}_l$, which can be obtained from $\mathcal{H}_l(\mathbb{R}^3)$ by restricting to the unit sphere. What is more is that the set $\{Y_m^l\}_{-l \leq m \leq l}$ is actually a basis for $\tilde{\mathcal{H}}_l$. In the case $l = 1$ this is clear: we have $\mathcal{H}_1(\mathbb{R}^3) = P_1(\mathbb{R}^3)$ and clearly $\{\frac{1}{\sqrt{2}}(x + iy), z, \frac{1}{\sqrt{2}}(x - iy)\}$ is a basis, and restricting this basis to the unit sphere gives the $l = 1$ spherical harmonics. For $l > 1$ proving

our claim requires a little more effort; see Problem 2.2. Another, simpler basis for $\mathcal{H}_1(\mathbb{R}^3)$ would be the *cartesian basis* $\{x, y, z\}$; physicists use the spherical harmonic basis because those functions are eigenfunctions of the orbital angular momentum operator L_z , which on $\mathcal{H}_l(\mathbb{R}^3)$ is represented⁸ by $L_z = -i(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x})$. We shall discuss the relationship between the two bases in detail later.

Not quite example $L^2([-a, a])$

From doing 1-D problems in quantum mechanics one already ‘knows’ that the set $\{e^{i \frac{n\pi x}{a}}\}_{n \in \mathbb{Z}}$ is a basis for $L^2([-a, a])$. There is a problem, however; we are used to taking infinite linear combinations of these basis vectors, but our definition above only allows for *finite* linear combinations. What is going on here? It turns out that $L^2([-a, a])$ has more structure than your average vector space: it is an infinite-dimensional *Hilbert space*, and for such spaces we have a generalized definition of a basis, one that allows for infinite linear combinations. We will discuss Hilbert spaces in Sect. 2.6

2.3 Components

One of the most useful things about introducing a basis for a vector space is that it allows us to write elements of the vector space as n -tuples, in the form of either column or row vectors, as follows: Given $v \in V$ and a basis $\mathcal{B} = \{e_i\}_{i=1, \dots, n}$ for V , we can write

$$v = \sum_{i=1}^n v^i e_i$$

for some numbers v^i , called the *components of v with respect to \mathcal{B}* . We can then represent v by the column vector, denoted $[v]_{\mathcal{B}}$, as

$$[v]_{\mathcal{B}} \equiv \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}$$

or the row vector

$$[v]_{\mathcal{B}}^T \equiv (v^1, v^2, \dots, v^n)$$

where the superscript T denotes the usual transpose of a vector. The subscript \mathcal{B} just reminds us which basis the components are referred to, and will be dropped if there is no ambiguity. With a choice of basis, then, every n -dimensional vector space

⁸As mentioned in the preface, the \hbar , which would normally appear in this expression, has been set to 1.

can be made to ‘look like’ \mathbb{R}^n or \mathbb{C}^n . Writing vectors in this way greatly facilitates computation, as we will see. One must keep in mind, however, that **vectors exist independently of any chosen basis**, and that their expressions as row or column vectors depend very much on the *choice* of basis \mathcal{B} .

To make this last point explicit, let us consider two bases for \mathbb{R}^3 : the standard basis

$$e_1 = (1, 0, 0)$$

$$e_2 = (0, 1, 0)$$

$$e_3 = (0, 0, 1)$$

and the alternate basis we introduced in Example 2.7

$$e'_1 = (1, 1, 0)$$

$$e'_2 = (0, 1, 1)$$

$$e'_3 = (1, 0, 1).$$

We will refer to these bases as \mathcal{B} and \mathcal{B}' , respectively. Let us consider the components of the vector e'_1 in both bases. If we expand in the basis \mathcal{B} , we have

$$e'_1 = 1 \cdot e_1 + 1 \cdot e_2 + 0 \cdot e_3$$

so

$$[e'_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad (2.10)$$

as expected. If we expand e'_1 in the basis \mathcal{B}' , however, we get

$$e'_1 = 1 \cdot e'_1 + 0 \cdot e'_2 + 0 \cdot e'_3$$

and so

$$[e'_1]_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

which is of course not the same as (2.10).⁹ For fun, we can also express the standard basis vector e_1 in the alternate basis \mathcal{B}' ; you can check that

$$e_1 = \frac{1}{2}e'_1 - \frac{1}{2}e'_2 + \frac{1}{2}e'_3$$

and so

⁹The simple form of $[e'_1]_{\mathcal{B}'}$ is no accident; you can easily check that if you express *any* set of basis vectors in the basis that they define, the resulting column vectors will just look like the standard basis.

$$[e_1]_{\mathcal{B}'} = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}.$$

The following examples will explore the notion of components in a few other contexts.

Example 2.11 *Rigid body motion*

One area of physics where the distinction between a vector and its expression as an ordered triple is crucial is rigid body motion. In this setting our vector space is \mathbb{R}^3 and we usually deal with two bases, an arbitrary but fixed *space axes* $K' = \{\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{z}}'\}$ and a time-dependent *body axes* $K = \{\hat{\mathbf{x}}(t), \hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)\}$ which is fixed relative to the rigid body. These are illustrated in Fig. 2.1. When we write down vectors in \mathbb{R}^3 , like the angular momentum vector L or the angular velocity vector ω , we must keep in mind what basis we are using, as the component expressions will differ drastically depending on the choice of basis. For example, if there is no external torque on a rigid body, $[L]_{K'}$ will be constant whereas $[L]_K$ will in general be time-dependent.

Example 2.12 *Different bases for \mathbb{C}^2*

As mentioned in Example 2.4, the vector space for a spin 1/2 particle is identifiable with \mathbb{C}^2 and the angular momentum operators are given by $L_i = \frac{1}{2}\sigma_i$. In particular, this means that

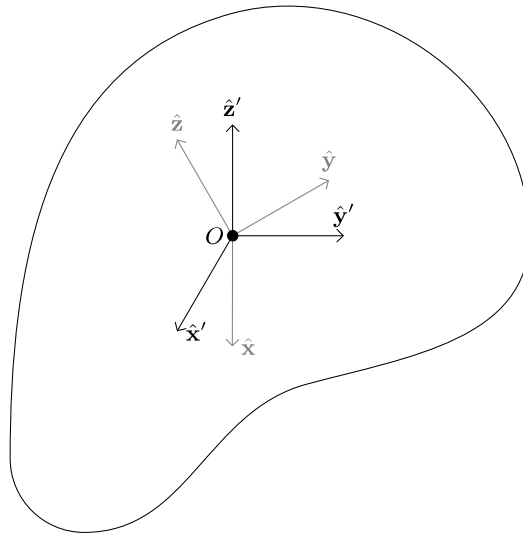


Fig. 2.1 Depiction of the fixed space axes K' and the time-dependent body axes K , in gray. K is attached to the rigid body and its basis vectors will change in time as the rigid body rotates

$$L_z = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

and so the standard basis vectors e_1, e_2 are eigenvectors of L_z with eigenvalues of $1/2$ and $-1/2$, respectively. Let us say, however, that we were interested in measuring L_x , where

$$L_x = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}.$$

Then we would need to use a basis \mathcal{B}' of L_x eigenvectors, which you can check are given by

$$\begin{aligned} e'_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ e'_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

If we are given the state e_1 and are asked to find the probability of measuring $L_x = 1/2$, then we need to expand e_1 in the basis \mathcal{B}' , which gives

$$e_1 = \frac{1}{\sqrt{2}}e'_1 + \frac{1}{\sqrt{2}}e'_2$$

and so

$$[e_1]_{\mathcal{B}'} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

This, of course, tells us that we have a probability of $1/2$ for measuring $L_x = +1/2$ (and the same for $L_x = -1/2$). Hopefully this convinces you that the distinction between a vector and its component representation is not just pedantic: it can be of real physical importance!

Example 2.13 $L^2([-a, a])$

We know from experience in quantum mechanics that all square integrable functions on an interval $[-a, a]$ have an expansion¹⁰

$$f = \sum_{m=-\infty}^{\infty} c_m e^{i \frac{m\pi x}{a}}$$

in terms of the ‘basis’ $\{\exp(i \frac{m\pi x}{a})\}_{m \in \mathbb{Z}}$. This expansion is known as the **Fourier series** of f , and we see that the c_m , commonly known as the **Fourier coefficients**, are nothing but the components of the vector f in the basis $\{e^{i \frac{m\pi x}{a}}\}_{m \in \mathbb{Z}}$.

¹⁰This fact is proved in most real analysis books, see Rudin [13].

2.4 Linear Operators

One of the basic notions in linear algebra, fundamental in quantum mechanics, is that of a linear operator. A **linear operator** on a vector space V is a function T from V to itself satisfying the **linearity** condition

$$T(cv + w) = cT(v) + T(w).$$

Sometimes we write Tv instead of $T(v)$. You should check that the set of all linear operators on V , with the obvious definitions of addition and scalar multiplication, forms a vector space, denoted $\mathcal{L}(V)$. You have doubtless seen many examples of linear operators: for instance, we can interpret a real $n \times n$ matrix as a linear operator on \mathbb{R}^n that acts on column vectors by matrix multiplication. Thus $M_n(\mathbb{R})$ (and, similarly, $M_n(\mathbb{C})$) can be viewed as vector spaces whose elements are themselves linear operators. In fact, that was exactly how we interpreted the vector subspace $H_2(\mathbb{C}) \subset M_2(\mathbb{C})$ in Example 2.4; in that case, we identified elements of $H_2(\mathbb{C})$ as the quantum-mechanical angular momentum operators. There are numerous other examples of quantum-mechanical linear operators: for instance, the familiar position and momentum operators \hat{x} and \hat{p} that act on $L^2([-a, a])$ by

$$\begin{aligned}(\hat{x}f)(x) &= xf(x) \\ (\hat{p}f)(x) &= -i \frac{\partial f}{\partial x}(x)\end{aligned}$$

as well as the angular momentum operators L_x , L_y , and L_z which act on $P_l(\mathbb{R}^3)$ by

$$\begin{aligned}L_x(f) &= -i \left(y \frac{\partial f}{\partial z} - z \frac{\partial f}{\partial y} \right) \\ L_y(f) &= -i \left(z \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial z} \right) \\ L_z(f) &= -i \left(x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x} \right).\end{aligned}\tag{2.11}$$

Another class of less familiar examples is given below.

Example 2.14 $\mathcal{L}(V)$ acting on $\mathcal{L}(V)$

We are familiar with linear operators taking vectors into vectors, but they can also be used to take linear operators into linear operators, as follows: Given $A, B \in \mathcal{L}(V)$, we can define a linear operator $\text{ad}_A \in \mathcal{L}(\mathcal{L}(V))$ acting on B by

$$\text{ad}_A(B) \equiv [A, B]$$

where $[\cdot, \cdot]$ indicates the commutator. This action of A on $\mathcal{L}(V)$ is called the **adjoint action** or **adjoint representation**. The adjoint representation has important applications in quantum mechanics; for instance, the Heisenberg picture emphasizes $\mathcal{L}(V)$ rather than V and interprets the Hamiltonian as an operator in the adjoint representation. In fact, for any observable A the Heisenberg equation of motion reads

$$\frac{dA}{dt} = i \operatorname{ad}_H(A). \quad (2.12)$$

□

One important property of a linear operator T is whether or not it is **invertible**, i.e. whether there exists a linear operator T^{-1} such that $TT^{-1} = T^{-1}T = I$ (where I is the identity operator).¹¹ You may recall that, in general, an inverse for a map F between two arbitrary sets A and B (not necessarily vector spaces!) exists if and only if F is both **one-to-one**, meaning

$$F(a_1) = F(a_2) \implies a_1 = a_2 \quad \forall a_1, a_2 \in A,$$

and **onto**, meaning that

$$\forall b \in B \text{ there exists } a \in A \text{ such that } F(a) = b.$$

If this is unfamiliar you should take a moment to convince yourself of this. In the particular case of a linear operator T on a vector space V (so that now instead of considering $F : A \rightarrow B$ we are considering $T : V \rightarrow V$, where T has the additional property of being linear), these two conditions are actually equivalent. Furthermore, they turn out also (see Exercise 2.7 below) to be equivalent to the statement

$$T(v) = 0 \implies v = 0,$$

so **T is invertible if and only if the only vector it sends to 0 is the zero vector.** The following exercises will help you parse these notions.

Exercise 2.6 Suppose V is finite-dimensional and let $T \in \mathcal{L}(V)$. Show that T being one-to-one is equivalent to T being onto. Feel free to introduce a basis to assist you in the proof.

Exercise 2.7 Suppose $T(v) = 0 \implies v = 0$. Show that this is equivalent to T being one-to-one, which by the previous exercise is equivalent to T being one-to-one *and* onto, which is then equivalent to T being invertible.

An important point to keep in mind is that **a linear operator is not the same thing as a matrix**; just as with vectors, the identification can only be made *once a basis is chosen*. For operators on finite-dimensional spaces this is done as follows: choose a basis $\mathcal{B} = \{e_i\}_{i=1, \dots, n}$. Then the action of T is determined by its action on the basis vectors:

$$T(v) = T\left(\sum_{i=1}^n v^i e_i\right) = \sum_{i=1}^n v^i T(e_i) = \sum_{i,j=1}^n v^i T_i^j e_j \quad (2.13)$$

where the numbers T_i^j , again called the **components of T with respect to \mathcal{B}** ,¹² are defined by

¹¹Throughout this text I will denote the identity operator or identity matrix; it will be clear from context which is meant.

¹²Nomenclature to be justified in the next chapter.

$$T(e_i) = \sum_{j=1}^n T_i^j e_j. \quad (2.14)$$

We then have

$$[v]_{\mathcal{B}} = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} \quad \text{and} \quad [T(v)]_{\mathcal{B}} = \begin{pmatrix} \sum_{i=1}^n v^i T_i^1 \\ \sum_{i=1}^n v^i T_i^2 \\ \vdots \\ \sum_{i=1}^n v^i T_i^n \end{pmatrix},$$

which looks suspiciously like matrix multiplication. In fact, we can define the **matrix of T in the basis \mathcal{B}** , denoted $[T]_{\mathcal{B}}$, by the matrix equation

$$[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}$$

where the product on the right-hand side is given by the usual matrix multiplication. Comparison of the last two equations then shows that

$$[T]_{\mathcal{B}} = \begin{pmatrix} T_1^1 & T_2^1 & \dots & T_n^1 \\ T_1^2 & T_2^2 & \dots & T_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ T_1^n & T_2^n & \dots & T_n^n \end{pmatrix}. \quad (2.15)$$

Thus, we really can use the components of T to represent it as a matrix, and once we do so the action of T becomes just matrix multiplication by $[T]_{\mathcal{B}}$! Furthermore, if we have two linear operators A and B and we define their **product** (or **composition**) AB as the linear operator

$$(AB)(v) \equiv A(B(v)),$$

you can then show that $[AB] = [A][B]$. Thus, composition of operators becomes matrix multiplication of the corresponding matrices.

Exercise 2.8 For two linear operators A and B on a vector space V , show that $[AB] = [A][B]$ in any basis.

Example 2.15 L_z , $\mathcal{H}_l(\mathbb{R}^3)$ and spherical harmonics

Recall that $\mathcal{H}_1(\mathbb{R}^3)$ is the set of all linear functions on \mathbb{R}^3 and that $\{rY_m^1\}_{-1 \leq m \leq 1} = \{\frac{1}{\sqrt{2}}(x + iy), z, \frac{1}{\sqrt{2}}(x - iy)\}$ and $\{x, y, z\}$ are both bases for this space.¹³ Now consider the familiar angular momentum operator $L_z = -i(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x})$ on this space. You can check that

¹³We have again ignored the overall normalization of the spherical harmonics to avoid unnecessary clutter.

$$L_z(x + iy) = x + iy$$

$$L_z(z) = 0$$

$$L_z(x - iy) = -(x - iy)$$

which implies by (2.14) that the components of L_z in the spherical harmonic basis are

$$(L_z)_1^1 = 1$$

$$(L_z)_1^2 = (L_z)_1^3 = 0$$

$$(L_z)_2^i = 0 \quad \forall i$$

$$(L_z)_3^3 = -1$$

$$(L_z)_3^1 = (L_z)_3^2 = 0.$$

Thus in the spherical harmonic basis,

$$[L_z]_{\{rY_m^l\}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

This of course just says that the wavefunctions $x + iy$, z and $x - iy$ have L_z eigenvalues of 1, 0, and -1 , respectively. Meanwhile,

$$L_z(x) = iy$$

$$L_z(y) = -ix$$

$$L_z(z) = 0$$

so in the cartesian basis,

$$[L_z]_{\{x,y,z\}} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.16)$$

a very differently looking matrix. Though the cartesian form of L_z is not usually used in physics texts, it has some very important mathematical properties, as we will see in Part II of this book. For a preview of these properties, see Problem 2.3.

Exercise 2.9 Compute the matrices of $L_x = -i(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y})$ and $L_y = -i(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z})$ acting on $\mathcal{H}_1(\mathbb{R}^3)$ in both the cartesian and spherical harmonic bases.

Before concluding this section we should remark that there is much more one can say about linear operators, particularly concerning eigenvectors, eigenvalues and diagonalization. Though these topics are relevant for physics, we will not need them in this text and good references for them abound, so we omit them. The interested reader can consult the first chapter of Sakurai [14] for a practical introduction, or Hoffman and Kunze [10] for a thorough discussion.

2.5 Dual Spaces

Another basic construction associated with a vector space, essential for understanding tensors and usually left out of the typical ‘mathematical methods for physicists’ courses, is that of a dual vector. Roughly speaking, a dual vector is an object that eats a vector and spits out a number. This may not sound like a very natural operation, but it turns out that this notion is what underlies bras in quantum mechanics, as well as the raising and lowering of indices in relativity. We will explore these applications in Sect. 2.7.

Now we give the precise definitions. Given a vector space V with scalars C , a **dual vector** (or **linear functional**) on V is a C -valued linear function f on V , where ‘linear’ again means

$$f(cv + w) = cf(v) + f(w). \quad (2.17)$$

(Note that $f(v)$ and $f(w)$ are scalars, so that on the left-hand side of (2.17) the addition takes place in V , whereas on the right side it takes place in C .) The set of all dual vectors on V is called the **dual space** of V , and is denoted V^* . It is easily checked that the usual definitions of addition and scalar multiplication and the zero function turn V^* into a vector space over C .

The most basic examples of dual vectors are the following: let $\{e_i\}$ be a (not necessarily finite) basis for V , so that an arbitrary vector v can be written as $v = \sum_{i=1}^n v^i e_i$. Then for each i , we can define a dual vector e^i by

$$\boxed{e^i(v) \equiv v^i}. \quad (2.18)$$

This just says that e^i ‘picks off’ the i th component of the vector v . Note that in order for this to make sense, a *basis has to be specified*.

A key property of dual vectors is that a dual vector f is entirely determined by its values on basis vectors: by linearity we have

$$\begin{aligned} f(v) &= f\left(\sum_{i=1}^n v^i e_i\right) \\ &= \sum_{i=1}^n v^i f(e_i) \\ &\equiv \sum_{i=1}^n v^i f_i \end{aligned} \quad (2.19)$$

where in the last line we have defined

$$\boxed{f_i \equiv f(e_i)}$$

which we unsurprisingly refer to as the **components** of f in the basis $\{e_i\}$. To justify this nomenclature, notice that the e^i defined in (2.18) satisfy

$$e^i(e_j) = \delta_j^i \quad (2.20)$$

where δ_j^i is the usual Kronecker delta. If V is finite-dimensional with dimension n , it is then easy to check (by evaluating both sides on basis vectors) that we can write¹⁴

$$f = \sum_{i=1}^n f_i e^i$$

so that the f_i really are the components of f . Since f was arbitrary, this means that the e^i span V^* . In Exercise 2.10 below you will show that the e^i are actually linearly independent, so $\{e^i\}_{i=1,\dots,n}$ is actually a *basis* for V^* . We sometimes say that the e^i are **dual to** the e_i . Note that we have shown that V and V^* always have the same dimension. We can use the dual basis $\{e^i\} \equiv \mathcal{B}^*$ to write f in components,

$$[f]_{\mathcal{B}^*} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

and in terms of the row vector $[f]_{\mathcal{B}^*}^T$ we can write (2.19) as

$$f(v) = [f]_{\mathcal{B}^*}^T [v]_{\mathcal{B}} = [f]^T [v]$$

where in the last equality we dropped the subscripts indicating the bases. Again, we allow ourselves to do this whenever there is no ambiguity about which basis for V we are using, and in all such cases we assume that the basis being used for V^* is just the one dual to the basis for V .

Finally, since $e^i(v) = v^i$ we note that we can alternatively think of the i th component of a vector as the value of e^i on that vector. This duality of viewpoint will crop up repeatedly in the rest of the text.

Exercise 2.10 By carefully working with the definitions, show that the e^i defined in (2.18) and satisfying (2.20) are linearly independent.

Example 2.16 Dual spaces of \mathbb{R}^n , \mathbb{C}^n , $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$

Consider the basis $\{e_i\}$ of \mathbb{R}^n and \mathbb{C}^n , where e_i is the vector with a 1 in the i th place and 0's everywhere else; this is just the basis described in Example 2.7. Now consider the element f^j of V^* which eats a vector in \mathbb{R}^n or \mathbb{C}^n and spits out the j th component; clearly $f^j(e_i) = \delta_i^j$ so the f^j are just the dual vectors e^j described above. Similarly, for $M_n(\mathbb{R})$ or $M_n(\mathbb{C})$ consider the dual vector f^{ij} defined by $f^{ij}(A) = A_{ij}$; these vectors are clearly dual to the E_{ij} and thus form the corresponding dual basis. While the f^{ij} may seem a little unnatural or artificial, you should note that there is one linear functional on $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ which is familiar: the **trace** functional, denoted Tr and defined by

¹⁴If V is infinite-dimensional then this may not work as the sum required may be infinite, and as mentioned before care must be taken in defining infinite linear combinations.

$$\text{Tr}(A) = \sum_{i=1}^n A_{ii}.$$

What are the components of Tr with respect to the f^{ij} ?

2.6 Non-degenerate Hermitian Forms

Non-degenerate Hermitian forms, of which the Euclidean dot product, Minkowski metric and Hermitian scalar product of quantum mechanics are but a few examples, are very familiar to most physicists. We introduce them here not just to formalize their definition but also to make the fundamental but usually unacknowledged connection between these objects and dual spaces.

A **non-degenerate Hermitian form** on a vector space V is a C -valued function $(\cdot | \cdot)$ which assigns to an ordered pair of vectors $v, w \in V$ a scalar, denoted $(v|w)$, having the following properties:

1. $(v|w_1 + cw_2) = (v|w_1) + c(v|w_2)$ (linearity in the second argument)
2. $(v|w) = \overline{(w|v)}$ (**Hermiticity**; the bar denotes complex conjugation)
3. For each $v \neq 0 \in V$, there exists $w \in V$ such that $(v|w) \neq 0$ (non-degeneracy)

Note that conditions 1 and 2 imply that $(cv|w) = \bar{c}(v|w)$, so $(\cdot | \cdot)$ is **conjugate-linear** in the first argument. Also note that for a real vector space, condition 2 implies that $(\cdot | \cdot)$ is **symmetric**, i.e. $(v|w) = (w|v)$ ¹⁵; in this case, $(\cdot | \cdot)$ is called a **metric**. Condition 3 is a little nonintuitive but will be essential in the connection with dual spaces. If, in addition to the above 3 conditions, the Hermitian form obeys

4. $(v|v) > 0$ for all $v \in V, v \neq 0$ (positive-definiteness)

then we say that $(\cdot | \cdot)$ is an **inner product**, and a vector space with such a Hermitian form is called an **inner product space**. In this case we can think of $(v|v)$ as the ‘length squared’ of the vector v , and the notation $\|v\| \equiv \sqrt{(v|v)}$ is sometimes used. Note that condition 4 implies 3. Our reason for separating condition 4 from the rest of the definition will become clear when we consider the examples. One very important use of non-degenerate Hermitian forms is to define preferred sets of bases known as **orthonormal bases**. Such bases $\mathcal{B} = \{e_i\}$ by definition satisfy $(e_i|e_j) = \pm\delta_{ij}$ and are extremely useful for computation, and ubiquitous in physics for that reason. If $(\cdot | \cdot)$ is positive-definite (hence an inner product), then orthonormal basis vectors satisfy $(e_i|e_j) = \delta_{ij}$ and may be constructed out of arbitrary bases by the Gram–Schmidt process. If $(\cdot | \cdot)$ is not positive-definite then orthonormal bases may still be constructed out of arbitrary bases, though the process is slightly more involved. See Hoffman and Kunze [10], Sects. 8.2 and 10.2 for details.

¹⁵In this case, $(\cdot | \cdot)$ is linear in the first argument as well as the second and would be referred to as **bilinear**.

Exercise 2.11 Let $(\cdot | \cdot)$ be an inner product. If a set of non-zero vectors e_1, \dots, e_k is *orthogonal*, i.e. $(e_i | e_j) = 0$ when $i \neq j$, show that they are linearly independent. Note that an *orthonormal* set (i.e. $(e_i | e_j) = \pm \delta_{ij}$) is just an orthogonal set in which the vectors have unit length.

Example 2.17 *The dot product (or Euclidean metric) on \mathbb{R}^n*

Let $v = (v^1, \dots, v^n)$, $w = (w^1, \dots, w^n) \in \mathbb{R}^n$. Define $(\cdot | \cdot)$ on \mathbb{R}^n by

$$(v|w) \equiv \sum_{i=1}^n v^i w^i.$$

This is sometimes written as $v \cdot w$. You should check that $(\cdot | \cdot)$ is an inner product, and that the standard basis given in Example 2.7 is an orthonormal basis.

Example 2.18 *The Hermitian scalar product on \mathbb{C}^n*

Let $v = (v^1, \dots, v^n)$, $w = (w^1, \dots, w^n) \in \mathbb{C}^n$. Define $(\cdot | \cdot)$ on \mathbb{C}^n by

$$(v|w) \equiv \sum_{i=1}^n \bar{v}^i w^i. \quad (2.21)$$

Again, you can check that $(\cdot | \cdot)$ is an inner product, and that the standard basis given in Example 2.7 is an orthonormal basis. Such inner products on complex vector spaces are sometimes referred to as *Hermitian scalar products* and are present on every quantum-mechanical vector space. In this example we see the importance of condition 2, manifested in the conjugation of the v^i in (2.21); if that conjugation was not there, a vector like $v = (i, 0, \dots, 0)$ would have $(v|v) = -1$ and $(\cdot | \cdot)$ would not be an inner product.

Exercise 2.12 Let $A, B \in M_n(\mathbb{C})$. Define $(\cdot | \cdot)$ on $M_n(\mathbb{C})$ by

$$(A|B) = \frac{1}{2} \text{Tr}(A^\dagger B). \quad (2.22)$$

Check that this is indeed an inner product. Also check that the basis $\{I, \sigma_x, \sigma_y, \sigma_z\}$ for $M_2(\mathbb{C})$ is orthonormal with respect to this inner product.

Example 2.19 *The Minkowski metric on 4-D spacetime*

Consider two vectors¹⁶ $v_i = (x_i, y_i, z_i, t_i) \in \mathbb{R}^4$, $i = 1, 2$. The Minkowski metric, denoted η , is defined to be¹⁷

$$\eta(v_1, v_2) \equiv x_1 x_2 + y_1 y_2 + z_1 z_2 - t_1 t_2. \quad (2.23)$$

¹⁶These are often called ‘events’ in the physics literature.

¹⁷We are, of course arbitrarily choosing the $+++-$ signature; we could equally well choose $---+$.

η is clearly linear in both its arguments (i.e. it is bilinear) and symmetric, hence satisfies conditions 1 and 2, and you will check condition 3 in Exercise 2.13 below. Notice that for $v = (1, 0, 0, 1)$, $\eta(v, v) = 0$ so η is *not* positive-definite, hence not an inner product. This is why we separated condition 4, and considered the more general non-degenerate Hermitian forms instead of just inner products.

Exercise 2.13 Let $v = (x, y, z, t)$ be an arbitrary non-zero vector in \mathbb{R}^4 . Show that η is non-degenerate by finding another vector w such that $\eta(v, w) \neq 0$.

We should point out here that the Minkowski metric can be written in components as a matrix, just as a linear operator can. Taking the standard basis $\mathcal{B} = \{e_i\}_{i=1,2,3,4}$ in \mathbb{R}^4 , we can define the **components of η** , denoted η_{ij} , as

$$\eta_{ij} \equiv \eta(e_i, e_j). \quad (2.24)$$

Then, just as was done for linear operators, you can check that if we define the **matrix of η in the basis \mathcal{B}** , denoted $[\eta]_{\mathcal{B}}$, as the matrix

$$[\eta]_{\mathcal{B}} = \begin{pmatrix} \eta_{11} & \eta_{21} & \eta_{31} & \eta_{41} \\ \eta_{12} & \eta_{22} & \eta_{32} & \eta_{42} \\ \eta_{13} & \eta_{23} & \eta_{33} & \eta_{43} \\ \eta_{14} & \eta_{24} & \eta_{34} & \eta_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (2.25)$$

then we can write

$$\eta(v_1, v_2) = [v_2]^T [\eta] [v_1] = (x_2, y_2, z_2, t_2) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ t_1 \end{pmatrix} \quad (2.26)$$

as some readers may be used to from computations in relativity. Note that the symmetry of η implies that $[\eta]_{\mathcal{B}}$ is a symmetric matrix for any basis \mathcal{B} .

Example 2.20 *The Hermitian scalar product on $L^2([-a, a])$*

For $f, g \in L^2([-a, a])$, define

$$(f|g) \equiv \frac{1}{2a} \int_{-a}^a \bar{f} g \, dx. \quad (2.27)$$

You can easily check that this defines an inner product on $L^2([-a, a])$, and that $\{e^{i \frac{n\pi x}{a}}\}_{n \in \mathbb{Z}}$ is an orthonormal set. What is more, this inner product turns $L^2([-a, a])$ into a **Hilbert Space**, which is an inner product space that is **complete**. The notion of completeness is a technical one, so we will not give its precise definition, but in the case of $L^2([-a, a])$ one can think of it as meaning roughly that a limit of a sequence of square-integrable functions is again square-integrable. Making this precise and proving it for $L^2([-a, a])$ is the subject of real analysis textbooks and far outside the scope of this text.¹⁸ We will thus content ourselves here with just mentioning

¹⁸See Rudin [13], for instance, for this and for proofs of all the claims made in this example.

completeness and noting that it is responsible for many of the nice features of Hilbert spaces, in particular the generalized notion of a basis which we now describe.

Given a Hilbert space \mathcal{H} and an orthonormal (and possibly infinite) set $\{e_i\} \subset \mathcal{H}$, the set $\{e_i\}$ is said to be an **orthonormal basis** for \mathcal{H} if

$$(e_i | f) = 0 \quad \forall i \implies f = 0. \quad (2.28)$$

You can check (see Exercise 2.14 below) that in the finite-dimensional case this definition is equivalent to our previous definition of an orthonormal basis. In the infinite-dimensional case, however, this definition differs substantially from the old one in that we no longer require $\text{Span}\{e_i\} = \mathcal{H}$ (recall that spans only include *finite* linear combinations). Does this mean, though, that we now allow *arbitrary* infinite combinations of the basis vectors? If not, which ones are allowed? For $L^2([-a, a])$, for which $\{e^{i\frac{n\pi x}{a}}\}_{n \in \mathbb{Z}}$ is an orthonormal basis, we mentioned in Example 2.13 that any $f \in L^2([-a, a])$ can be written as

$$f = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi x}{a}} \quad (2.29)$$

where

$$\frac{1}{2a} \int_{-a}^a |f|^2 dx = \sum_{n=-\infty}^{\infty} |c_n|^2 < \infty. \quad (2.30)$$

(The first equality in (2.30) should be familiar from quantum mechanics and follows from Exercise 2.15 below.) The converse to this is also true, and this is where the completeness of $L^2([-a, a])$ is essential: if a set of numbers c_n satisfy (2.30), then the series

$$g(x) \equiv \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi x}{a}} \quad (2.31)$$

converges, yielding a square-integrable function g . So $L^2([-a, a])$ is the set of all expressions of the form (2.29), *subject to the condition* (2.30). Now we know how to think about infinite-dimensional Hilbert spaces and their bases: a basis for a Hilbert space is an infinite set whose infinite linear combinations, *together with some suitable convergence condition*, form the entire vector space.

Exercise 2.14 Show that the definition (2.28) of a Hilbert space basis is equivalent to our original definition of a basis for a *finite*-dimensional inner product space V .

Exercise 2.15 Show that for $f = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi x}{a}}$, $g = \sum_{m=-\infty}^{\infty} d_m e^{i\frac{m\pi x}{a}} \in L^2([-a, a])$,

$$(f | g) = \sum_{n=-\infty}^{\infty} \bar{c}_n d_n \quad (2.32)$$

so that $(\cdot | \cdot)$ on $L^2([-a, a])$ can be viewed as the infinite-dimensional version of (2.21), the standard Hermitian scalar product on \mathbb{C}^n .

2.7 Non-degenerate Hermitian Forms and Dual Spaces

We are now ready to explore the connection between dual vectors and non-degenerate Hermitian forms. Given a non-degenerate Hermitian form $(\cdot | \cdot)$ on a finite-dimensional vector space V , we can associate to any $v \in V$ a dual vector $\tilde{v} \in V^*$ defined by

$$\tilde{v}(w) \equiv (v|w). \quad (2.33)$$

This defines a very important map,

$$\boxed{\begin{array}{l} L : V \rightarrow V^* \\ v \mapsto \tilde{v} \end{array}}$$

which will crop up repeatedly in this chapter and the next. We will sometimes write \tilde{v} as $L(v)$ or $v|\cdot$, and refer to it as the **metric dual** of v .¹⁹

Now, L is conjugate-linear since for $v = cx + z$, $v, z, x \in V$,

$$\tilde{v}(w) = (v|w) = (cx + z|w) = \bar{c}(x|w) + (z|w) = \bar{c}\tilde{x}(w) + \tilde{z}(w) \quad (2.34)$$

so

$$\tilde{v} = L(cx + z) = \bar{c}\tilde{x} + \tilde{z} = \bar{c}L(x) + L(z). \quad (2.35)$$

In Exercise 2.16 below you will show that the non-degeneracy of $(\cdot | \cdot)$ implies that L is one-to-one and onto, so L is an invertible map from V to V^* . This allows us to identify V with V^* , a fact we will discuss further below.

Before we get to some examples, it is worth pointing out one potential point of confusion: if we have a basis $\{e_i\}_{i=1,\dots,n}$ and corresponding dual basis $\{e^i\}_{i=1,\dots,n}$, **it is not necessarily true that a dual vector e^i in the dual basis is the same as the metric dual $L(e_i)$** ; in particular, the dual basis vector e^i is defined relative to a whole *basis* $\{e_i\}_{i=1,\dots,n}$, whereas the metric dual $L(e_i)$ only depends on what e_i is, and does not care if we change the other basis vectors. Furthermore, $L(e_i)$ depends on your non-degenerate Hermitian form (that is why we call it a *metric dual*), whereas e^i does not (remember that we introduced dual basis vectors in Sect. 2.5, before we even knew what non-degenerate Hermitian forms were!). You may wonder if there are special circumstances when we *do* have $e^i = L(e_i)$; this is Exercise 2.17.

With that caveat let us proceed to some examples, where you will see that you are already familiar with L from a couple of different contexts.

Exercise 2.16 Use the non-degeneracy of $(\cdot | \cdot)$ to show that L is one-to-one, i.e. that $L(v) = L(w) \implies v = w$. Combine this with the argument used in Exercise 2.7 to show that L is onto as well.

Exercise 2.17 Given a basis $\{e_i\}_{i=1,\dots,n}$, under what circumstances do we have $e^i = \tilde{e}_i$ for all i ?

¹⁹The \cdot in the notation $(v|\cdot)$ signifies the slot into which a vector w is to be inserted, yielding the number $(v|w)$.

Example 2.21 Bras and kets in quantum mechanics

Let \mathcal{H} be a quantum-mechanical Hilbert space with inner product $(\cdot | \cdot)$. In Dirac notation, a vector $\psi \in \mathcal{H}$ is written as a ket $|\psi\rangle$ and the inner product (ψ, ϕ) is written $\langle\psi|\phi\rangle$. What about bras, written as $\langle\psi|$? What, exactly, are they? Most quantum mechanics texts gloss over their definition, just telling us that they are in 1–1 correspondence with kets and can be combined with kets as $\langle\psi|\phi\rangle$ to get a scalar. We are also told that the correspondence between bras and kets is conjugate-linear, i.e. that the bra corresponding to $c|\psi\rangle$ is $\bar{c}\langle\psi|$. From what we have seen in this section, it is now clear that **bras really are dual vectors**, labeled in the same way as regular vectors, because the map L allows us to identify the two. In short, $\langle\psi|$ is really just $L(\psi)$, or equivalently $(\psi|\cdot)$.

Example 2.22 Raising and lowering indices in relativity

Consider \mathbb{R}^4 with the Minkowski metric, let $\mathcal{B} = \{e_\mu\}_{\mu=1,\dots,4}$ and $\mathcal{B}' = \{e^\mu\}_{\mu=1,\dots,4}$ be the standard basis and dual basis for \mathbb{R}^4 (we use a Greek index to conform with standard physics notation²⁰), and let $v = \sum_{\mu=1}^4 v^\mu e_\mu \in \mathbb{R}^4$. What are the components of the dual vector \tilde{v} in terms of the v^μ ? Well, as we saw in Sect. 2.5, the components of a dual vector are just given by evaluation on the basis vectors, so

$$\tilde{v}_\mu = \tilde{v}(e_\mu) = (v|e_\mu) = \sum_\nu v^\nu (e_\nu|e_\mu) = \sum_\nu v^\nu \eta_{\nu\mu}. \quad (2.36)$$

In matrices, this reads

$$[\tilde{v}]_{\mathcal{B}'} = [\eta]_{\mathcal{B}}[v]_{\mathcal{B}} \quad (2.37)$$

so matrix multiplication of a vector by the metric matrix gives the corresponding dual vector in the dual basis. Thus, the map L is implemented in coordinates by $[\eta]$. Now, we mentioned above that L is invertible; what does L^{-1} look like in coordinates? Well, by the above, L^{-1} should be given by matrix multiplication by $[\eta]^{-1}$, the matrix inverse to $[\eta]$. Denoting the components of this matrix by $\eta^{\mu\nu}$ (so that $\eta^{\tau\mu}\eta_{\nu\tau} = \delta_\nu^\mu$) and writing $\tilde{f} \equiv L^{-1}(f)$ where f is a dual vector, we have

$$[\tilde{f}] = [\eta]^{-1}[f] \quad (2.38)$$

or in components

$$\tilde{f}^\mu = \sum_\nu \eta^{\mu\nu} f_\nu. \quad (2.39)$$

The expressions in (2.36) and (2.39) are probably familiar to you. In physics one usually works with components of vectors, and in relativity the numbers v^μ

²⁰As long as we are talking about ‘standard’ physics notation, you should also be aware that in many texts the indices run from 0 to 3 instead of 1 to 4, and in that case the zeroth coordinate corresponds to time.

are called the **contravariant components** of v and the numbers $v_\mu \equiv \sum_\nu v^\nu \eta_{\nu\mu}$ of (2.36) are referred to as the **covariant components** of v . We see now that the contravariant components of a vector are just its usual components, **while its covariant components are actually the components of the associated dual vector** \tilde{v} . For a dual vector f , the situation is reversed—the covariant components f_μ are its actual components, and the contravariant components are the components of \tilde{f} . Since L allows us to turn vectors into dual vectors and vice versa, we usually do not bother trying to figure out whether something is ‘really’ a vector or a dual vector; it can be either, depending on which components we use.

The above discussion shows that the familiar process of “raising” and “lowering” indices is just the application of the map L (and its inverse) in components. For an interpretation of $[\eta]^{-1}$ as the matrix of a metric on \mathbb{R}^{4*} , see Problem 2.7.

Exercise 2.18 Consider \mathbb{R}^3 with the Euclidean metric. Show that the covariant and contravariant components of a vector in an orthonormal basis are identical. This explains why we never bother with this terminology, nor the concept of dual spaces, in basic physics where \mathbb{R}^3 is the relevant vector space. Is the same true for \mathbb{R}^4 with the Minkowski metric?

Example 2.23 $L^2([-a, a])$ and its dual

In our above discussion of the map L we stipulated that V should be finite-dimensional. Why? If you examine the discussion closely, you will see that the only place where we use the finite-dimensionality of V is in showing that L is onto. Does this mean that L is not necessarily onto in infinite dimensions? Consider the Dirac Delta functional $\delta \in L^2([-a, a])^*$. Does

$$\delta(g) = g(0) \stackrel{?}{=} (\delta(x)|g) \quad (2.40)$$

for some function $\delta(x)$? If we write g as $g(x) = \sum_{n=-\infty}^{\infty} d_n e^{i \frac{n\pi x}{a}}$, then simply evaluating g at $x = 0$ gives

$$g(0) = \sum_{n=-\infty}^{\infty} d_n \stackrel{?}{=} (\delta(x)|g) \quad (2.41)$$

which, when compared with (2.32), tells us that the function $\delta(x)$ must have Fourier coefficients $c_n = 1$ for all n . Such c_n , however, do not satisfy (2.30) and hence $\delta(x)$ cannot be a square-integrable function. So the dual vector δ is not in the image of the map L , hence L is not onto in the case of $L^2([-a, a])$.

2.8 Problems

2.1 Prove that $L^2([-a, a])$ is closed under addition. You will need the triangle inequality, as well as the following inequality, valid for all $\lambda \in \mathbb{R}$: $0 \leq \int_{-a}^a (|f| + \lambda|g|)^2 dx$.

2.2 In this problem we show that $\{r^l Y_m^l\}$ is a basis for $\mathcal{H}_l(\mathbb{R}^3)$, which implies that $\{Y_m^l\}$ is a basis for $\tilde{\mathcal{H}}_l$. We will gloss over a few subtleties here; for a totally rigorous discussion see Sternberg [16] or our discussion in Chap. 4.

(a) Let $f \in \mathcal{H}_l(\mathbb{R}^3)$, and write f as $f = r^l Y(\theta, \phi)$. Then we know that

$$\Delta_{S^2} Y = -l(l+1)Y. \quad (2.42)$$

If you have never done so, use the expression (2.4) for Δ_{S^2} and the expressions (2.11) for the angular momentum operators to show that

$$\Delta_{S^2} = L_x^2 + L_y^2 + L_z^2 \equiv \mathbf{L}^2$$

so that (2.42) says that Y is an eigenfunction of \mathbf{L}^2 , as expected. You will need to convert between cartesian and spherical coordinates. The theory of angular momentum²¹ then tells us that $\mathcal{H}_l(\mathbb{R}^3)$ has dimension $2l+1$.

(b) Exhibit a basis for $\mathcal{H}_l(\mathbb{R}^3)$ by considering the function $f_0^l \equiv (x+iy)^l$ and showing that

$$L_z(f_0^l) = lf_0^l, \quad L_+(f_0^l) \equiv (L_x + iL_y)(f_0^l) = 0.$$

The theory of angular momentum then tells us that $(L_-)^k f_0^l \equiv f_k^l$ satisfies $L_z(f_k^l) = (l-k)f_k^l$ and that $\{f_k^l\}_{0 \leq k \leq 2l}$ is a basis for $\mathcal{H}_l(\mathbb{R}^3)$.

(c) Writing $f_k^l = r^l Y_{l-k}^l$ we see that Y_m^l satisfies $\mathbf{L}^2 Y_m^l = -l(l+1)Y_m^l$ and $L_z Y_m^l = mY_m^l$ as expected. Now use this definition of Y_m^l to compute all the spherical harmonics for $l=1, 2$ and show that this agrees, up to normalization, with the spherical harmonics as tabulated in any quantum mechanics textbook. If you read Example 2.6 and did Exercise 2.5 then all you have to do is compute f_k^1 , $0 \leq k \leq 2$ and f_k^2 , $0 \leq k \leq 4$ and show that these functions agree with the ones given there.

2.3 In discussions of quantum mechanics you may have heard the phrase “angular momentum generates rotations”. What this means is that if one takes a component of the angular momentum such as L_z and exponentiates it, i.e. if one considers the operator

$$\begin{aligned} \exp(-i\phi L_z) &\equiv \sum_{n=0}^{\infty} \frac{1}{n!} (-i\phi L_z)^n \\ &= I - i\phi L_z + \frac{1}{2!} (-i\phi L_z)^2 + \frac{1}{3!} (-i\phi L_z)^3 + \dots \end{aligned}$$

(the usual power series expansion for e^x) then one gets the operator which represents a rotation about the z axis by an angle ϕ . Confirm this in one instance by explicitly summing this power series for the operator $[L_z]_{\{x,y,z\}}$ of Example 2.15 to get

²¹ See Sakurai [14] or Gasiorowicz [5] or our discussion in Chap. 4.

$$\exp(-i\phi[L_z]_{\{x,y,z\}}) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the usual matrix for a rotation about the z -axis.

2.4 Let V be finite-dimensional, and consider the ‘double dual’ space $(V^*)^*$. Define a map

$$J : V \rightarrow (V^*)^* \\ v \mapsto J_v$$

where J_v acts on $f \in V^*$ by

$$J_v(f) \equiv f(v).$$

Show that J is linear, one-to-one, and onto. This means that we can identify $(V^*)^*$ with V itself, and so the process of taking duals repeats itself after two iterations. Note that no non-degenerate Hermitian form was involved in the definition of J !

2.5 Consider a linear operator A on a vector space V . We can define a linear operator on V^* called the **transpose** of A and denoted by A^T as follows:

$$(A^T(f))(v) \equiv f(Av) \quad \text{where } v \in V, f \in V^*.$$

If \mathcal{B} is a basis for V and \mathcal{B}^* the corresponding dual basis, show that

$$[A^T]_{\mathcal{B}^*} = [A]_{\mathcal{B}}^T.$$

Thus the transpose of a matrix really has meaning; it is the matrix representation of the transpose of the linear operator represented by the original matrix!

2.6 This problem will explore the notion of the Hermitian adjoint of a linear operator.

- (a) Let A be a linear operator on a finite-dimensional, real or complex vector space V with inner product $(\cdot | \cdot)$. Using the transpose A^T from the previous problem, as well as the map $L : V \rightarrow V^*$ defined in Sect. 2.7, we can construct a *new* linear operator A^\dagger ; this is known as the **Hermitian adjoint** of A , and is defined as

$$A^\dagger \equiv L^{-1} \circ A^T \circ L : V \rightarrow V. \quad (2.43)$$

Show that A^\dagger satisfies

$$(A^\dagger v | w) = (v | Aw), \quad (2.44)$$

which is the equation that is usually taken as the definition of the adjoint operator. The advantage of our definition (2.43) is that it gives an interpretation to A^\dagger ; it is the operator you get by transplanting A^T , which originally acts on V^* , over to V via L .

- (b) Show that in an orthonormal basis $\{e_i\}_{i=1,\dots,n}$,

$$[A^\dagger] = [A]^\dagger,$$

where the dagger outside the brackets denotes the usual conjugate transpose of a matrix (if V is a real vector space, then the dagger outside the brackets will reduce to just the transpose). You may want to prove and use the fact $A_j^i = e^i(Ae_j)$.

- (c) If A satisfies $A = A^\dagger$, A is then said to be **self-adjoint** or **Hermitian**. Since A^\dagger is defined with respect to an inner product, self-adjointness indicates a certain compatibility between A and the inner product. Show that even when V is a complex vector space, any eigenvalue of A must be real if A is self-adjoint.
- (d) In part (b) you showed that *in an orthonormal basis* the matrix of a Hermitian operator is a Hermitian matrix. Is this necessarily true in a non-orthonormal basis?

2.7 Let g be a non-degenerate bilinear form on a vector space V (we have in mind the Euclidean metric on \mathbb{R}^3 or the Minkowski metric on \mathbb{R}^4). Pick an arbitrary (not necessarily orthonormal) basis, let $[g]^{-1}$ be the matrix inverse of $[g]$ in this basis, and write $g^{\mu\nu}$ for the components of $[g]^{-1}$. Also let $f, h \in V^*$. Define a non-degenerate bilinear form \tilde{g} on V^* by

$$\tilde{g}(f, h) \equiv g(\tilde{f}, \tilde{h}) \quad (2.45)$$

where $\tilde{f} = L^{-1}(f)$ as in Example 2.22. Show that

$$\tilde{g}^{\mu\nu} \equiv \tilde{g}(e^\mu, e^\nu) = g^{\mu\nu} \quad (2.46)$$

so that $[g]^{-1}$ is truly a matrix representation of a non-degenerate bilinear form on V^* .

2.8 In this problem we will acquaint ourselves with $P(\mathbb{R})$, the set of polynomials in one variable x with real coefficients. We will also meet several bases for this space which you should find familiar.

- (a) $P(\mathbb{R})$ is the set of all functions of the form

$$f(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n \quad (2.47)$$

where n is arbitrary. Verify that $P(\mathbb{R})$ is a (real) vector space. Then show that $P(\mathbb{R})$ is infinite-dimensional by showing that, for any finite set $S \subset P(\mathbb{R})$, there is a polynomial that is not in $\text{Span } S$. Exhibit a simple infinite basis for $P(\mathbb{R})$.

- (b) Compute the matrix corresponding to the operator $\frac{d}{dx} \in \mathcal{L}(P(\mathbb{R}))$ with respect to the basis you found in part (a).
- (c) One can turn $P(\mathbb{R})$ into an inner product space by considering inner products of the form

$$(f|g) \equiv \int_a^b f(x)g(x)W(x)dx \quad (2.48)$$

where $W(x)$ is a nonnegative **weight function**. One can then take the basis $\mathcal{B} = \{1, x, x^2, x^3, \dots\}$ and apply the Gram–Schmidt process to get an orthogonal basis. With the proper choice of range of integration $[a, b]$ and weight function $W(x)$, we can obtain (up to normalization) the various orthogonal polynomials one meets in studying the various differential equations that arise in electrostatics and quantum mechanics.

- (i) Let $[a, b] = [-1, 1]$ and $W(x) = 1$. Consider the set $S = \{1, x, x^2, x^3\} \subset \mathcal{B}$. Apply the Gram–Schmidt process to this set to get (up to normalization) the first four **Legendre Polynomials**

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x).$$

The Legendre Polynomials show up in the solutions to the differential equation (2.42), where we make the identification $x = \cos \theta$. Since $-1 \leq \cos \theta \leq 1$, this explains the range of integration in the inner product.

- (ii) Now let $[a, b] = (-\infty, \infty)$ and $W(x) = e^{-x^2}$. Use Gram–Schmidt on S to get the first four **Hermite Polynomials**

$$H_0(x) = 1$$

$$H_1(x) = 2x$$

$$H_2(x) = 4x^2 - 2$$

$$H_3(x) = 8x^3 - 12x.$$

These polynomials arise in the solution to the Schrödinger equation for a one-dimensional harmonic oscillator. Note that the range of integration corresponds to the range of the position variable, as expected.

- (iii) Finally, let $[a, b] = (0, \infty)$ and $W(x) = e^{-x}$. Again, use Gram–Schmidt on S to get the first four **Laguerre Polynomials**

$$L_0(x) = 1$$

$$L_1(x) = -x + 1$$

$$L_2(x) = \frac{1}{2}(x^2 - 4x + 2)$$

$$L_3(x) = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6).$$

These polynomials arise as solutions to the radial part of the Schrödinger equation for the hydrogen atom. In this case x is interpreted as a radial variable, hence the range of integration $(0, \infty)$.

Chapter 3

Tensors

Now that we are familiar with vector spaces we can finally approach our main subject, tensors. We will give the modern component-free definition, from which will follow the usual transformation laws that used to *be* the definition.

From here on out we will employ the ***Einstein summation convention***, which is that whenever an index is repeated in an expression, once as a superscript and once as a subscript, then summation over that index is implied. Thus an expression like $v = \sum_{i=1}^n v^i e_i$ becomes $v = v^i e_i$. We will comment on this convention in Sect. 3.2. Also, in this chapter we will treat mostly finite-dimensional vector spaces and ignore the complications and technicalities that arise in the infinite-dimensional case. In the few examples where we apply our results to infinite-dimensional spaces, you should rest assured that these applications are legitimate (if not explicitly justified), and that rigorous justification can be found in the literature.

3.1 Definition and Examples

A **tensor of type** (r, s) on a vector space V is a C -valued function T on

$$\underbrace{V \times \cdots \times V}_{r \text{ times}} \times \underbrace{V^* \times \cdots \times V^*}_{s \text{ times}} \quad (3.1)$$

which is linear in each argument, i.e.

$$\begin{aligned} T(v_1 + cw, v_2, \dots, v_r, f_1, \dots, f_s) \\ = T(v_1, \dots, v_r, f_1, \dots, f_s) + cT(w, v_2, \dots, f_1, \dots, f_s) \end{aligned} \quad (3.2)$$

and similarly for all the other arguments. This property is called **multilinearity**. Note that dual vectors are $(1, 0)$ tensors, and that vectors can be viewed as $(0, 1)$ tensors:

$$v(f) \equiv f(v) \quad \text{where } v \in V, f \in V^*. \quad (3.3)$$

Similarly, linear operators can be viewed as $(1, 1)$ tensors as

$$A(v, f) \equiv f(Av). \quad (3.4)$$

We take $(0, 0)$ tensors to be scalars, as a matter of convention. You will show in Exercise 3.1 below that the set of all tensors of type (r, s) on a vector space V , denoted $\mathcal{T}_s^r(V)$ or just \mathcal{T}_s^r , form a vector space. This should not come as much of a surprise since we already know that vectors, dual vectors and linear operators all form vector spaces. Also, just as linearity implies that dual vectors and linear operators are determined by their values on basis vectors, multilinearity implies the same thing for general tensors. To see this, let $\{e_i\}_{i=1,\dots,n}$ be a basis for V and $\{e^i\}_{i=1,\dots,n}$ the corresponding dual basis. Then, denoting the i th component of the vector v_p as v_p^i and the j th component of the dual vector f_q as f_{qj} , we have (by repeated application of multilinearity)

$$\begin{aligned} T(v_1, \dots, v_r, f_1, \dots, f_s) &= v_1^{i_1} \dots v_r^{i_r} f_{1j_1} \dots f_{sj_s} T(e_{i_1}, \dots, e_{i_r}, e^{j_1}, \dots, e^{j_s}) \\ &\equiv v_1^{i_1} \dots v_r^{i_r} f_{1j_1} \dots f_{sj_s} T_{i_1, \dots, i_r}^{j_1, \dots, j_s} \end{aligned}$$

where, as before, the numbers

$$T_{i_1 \dots i_r}^{j_1 \dots j_s} \equiv T(e_{i_1}, \dots, e_{i_r}, e^{j_1}, \dots, e^{j_s}) \quad (3.5)$$

are called the **components of T** in the basis $\{e_i\}_{i=1,\dots,n}$. You should check that this definition of the components of a tensor, when applied to vectors, dual vectors, and linear operators, agrees with the definitions given earlier. Also note that (3.5) gives us a concrete way to think about the components of tensors: they are the *values of the tensor on the basis vectors*.

Exercise 3.1 By choosing suitable definitions of addition and scalar multiplication, show that $\mathcal{T}_s^r(V)$ is a vector space.

If we have a non-degenerate bilinear form on V , then we may change the type of T by precomposing with the map L or L^{-1} . If T is of type $(1, 1)$ with components T_i^j , for instance, then we may turn it into a tensor \tilde{T} of type $(2, 0)$ by defining $\tilde{T}(v, w) = T(v, L(w))$. This corresponds to lowering the second index, and we write the components of \tilde{T} as T_{ij} , omitting the tilde since the fact that we lowered the second index implies that we precomposed with L . This is in accord with the conventions in relativity, where given a vector $v \in \mathbb{R}^4$ we write v_μ for the components of \tilde{v} when we should really write \tilde{v}_μ . From this point on, if we have a non-degenerate bilinear form on a vector space then we permit ourselves to raise and lower indices at will and without comment. In such a situation we often do not discuss the type of a tensor, speaking instead of its **rank**, equal to $r + s$, which obviously does not change as we raise and lower indices.

Example 3.1 Linear operators in quantum mechanics

Thinking about linear operators as $(1, 1)$ tensors may seem a bit strange, but in fact this is what one does in quantum mechanics all the time! Given an operator H on a quantum mechanical Hilbert space spanned by orthonormal vectors $\{e_i\}$ (which in Dirac notation we would write as $\{|i\rangle\}$), we usually write $H|i\rangle$ for $H(e_i)$, $\langle j|i\rangle$ for $\tilde{e}_j(e_i) = (e_j|e_i)$, and $\langle j|H|i\rangle$ for $(e_j|He_i)$. Thus, (3.4) would tell us that (using orthonormal basis vectors instead of arbitrary vectors)

$$\begin{aligned}
H_i^j &= H(e_i, e^j) \\
&= e^j(He_i) \\
&= \langle j | H | i \rangle
\end{aligned}$$

where we converted to Dirac notation in the last equality to obtain the familiar quantum mechanical expression for the components of a linear operator. These components are often referred to as **matrix elements**, since when we write operators as matrices the elements of the matrices are just the components arranged in a particular fashion, as in (2.15).

Example 3.2 The Levi-Civita tensor

Consider the $(3, 0)$ **Levi-Civita tensor** ϵ on \mathbb{R}^3 defined by

$$\epsilon(u, v, w) \equiv (u \times v) \cdot w, \quad u, v, w \in \mathbb{R}^3. \quad (3.6)$$

You will check below that ϵ really is multilinear, hence a tensor. It is well-known from vector calculus that $(u \times v) \cdot w$ is the (oriented) volume of a parallelepiped spanned by u , v , and w (see Fig. 3.1), so one can think of the Levi-Civita tensor as a kind of ‘volume operator’ which eats three vectors and spits out the volume that they span.

What about the components of the Levi-Civita tensor? If $\{e_1, e_2, e_3\}$ is the standard basis for \mathbb{R}^3 then (3.6) yields

$$\begin{aligned}
\epsilon_{ijk} &= \epsilon(e_i, e_j, e_k) \\
&= (e_i \times e_j) \cdot e_k \\
&= \bar{\epsilon}_{ijk}
\end{aligned}$$

where $\bar{\epsilon}_{ijk}$ is the usual Levi-Civita symbol (defined below). Thus the Levi-Civita symbol represents the components of an actual tensor, the Levi-Civita tensor! Furthermore, keeping in mind the interpretation of the Levi-Civita tensor as a volume operator, as well as the fact that the components ϵ_{ijk} are just the values of the tensor on the basis vectors, then we find that the usual definition of the Levi-Civita symbol,

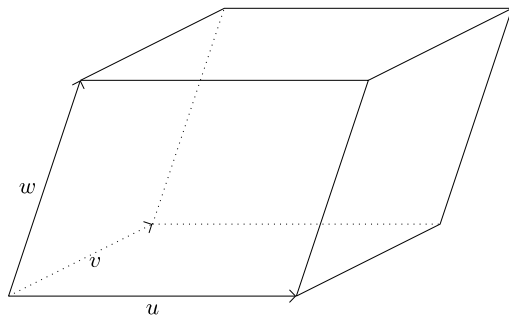


Fig. 3.1 The parallelepiped spanned by u , v and w

$$\bar{\epsilon}_{ijk} = \begin{cases} +1 & \text{if } \{ijk\} = \{1, 2, 3\}, \{2, 3, 1\}, \text{ or } \{3, 1, 2\} \\ -1 & \text{if } \{ijk\} = \{3, 2, 1\}, \{1, 3, 2\}, \text{ or } \{2, 1, 3\} \\ 0 & \text{otherwise,} \end{cases}$$

is just telling us, for instance, that a parallelepiped spanned by $\{e_1, e_2, e_3\}$ has oriented volume $+1$!

Exercise 3.2 Verify that the Levi-Civita tensor as defined by (3.6) really is multilinear.

Example 3.3 The moment of inertia tensor

The moment of inertia tensor, denoted \mathcal{I} , is the symmetric $(2, 0)$ tensor on \mathbb{R}^3 which, when evaluated on the angular velocity vector, yields the kinetic energy of a rigid body, i.e.

$$\frac{1}{2} \mathcal{I}(\omega, \omega) = KE. \quad (3.7)$$

Alternatively we can raise an index on \mathcal{I} and define it to be the linear operator which eats the angular velocity and spits out the angular momentum, i.e.

$$\mathbf{L} = \mathcal{I}\omega. \quad (3.8)$$

Equations (3.7) and (3.8) are most often seen in components (referred to a cartesian basis), where they read

$$KE = [\omega]^T [\mathcal{I}][\omega] \\ [\mathbf{L}] = [\mathcal{I}][\omega].$$

Note that since we raise and lower indices with an inner product and usually use orthonormal bases, the components of \mathcal{I} when viewed as a $(2, 0)$ tensor and when viewed as a $(1, 1)$ tensor are the same, cf. Exercise 2.18. We'll discuss \mathcal{I} further in Sect. 3.6.

Example 3.4 Multipole moments

It is a standard result from electrostatics that the scalar potential $\Phi(\mathbf{r})$ of a charge distribution $\rho(\mathbf{r}')$ localized around the origin in \mathbb{R}^3 can be expanded in a Taylor series as¹

$$\Phi(\mathbf{r}) = \frac{1}{4\pi} \left[\frac{Q_0}{r} + \frac{Q_1(\mathbf{r})}{r^3} + \frac{1}{2!} \frac{Q_2(\mathbf{r}, \mathbf{r})}{r^5} + \frac{1}{3!} \frac{Q_3(\mathbf{r}, \mathbf{r}, \mathbf{r})}{r^7} + \dots \right] \quad (3.9)$$

where the Q_r are r th rank tensors known as the **multipole moments** of the charge distribution $\rho(\mathbf{r}')$. The first few multipole moments are familiar to most physicists: the first, Q_0 , is just the total charge or **monopole moment** of the charge distribution and is given by

¹Here and below we set all physical constants such as c and ϵ_0 equal to 1.

$$Q_0 = \int \rho(\mathbf{r}') d^3 r'. \quad (3.10)$$

The second, Q_1 , is a dual vector known as the **dipole moment** (often denoted as \mathbf{p}), which has components

$$p_i = \int x'_i \rho(\mathbf{r}') d^3 r'. \quad (3.11)$$

The third multipole moment, Q_2 , is known as the **quadrupole moment** and has components given by

$$Q_{ij} = \int (3x'_i x'_j - r'^2 \delta_{ij}) d^3 r'. \quad (3.12)$$

Notice that the Q_{ij} are symmetric in i and j , and that $\sum_i Q_{ii} = 0$. Analogous properties hold for the higher order multipole moments as well (i.e. the **octopole moment** Q_3 has components Q_{ijk} which are totally symmetric and which satisfy $\sum_i Q_{iij} = \sum_i Q_{iji} = \sum_i Q_{jii} = 0$). We will explain these curious features of the Q_r at the end of this chapter.

Example 3.5 Metric tensors

We met the Euclidean metric on \mathbb{R}^n in Example 2.17 and the Minkowski metric on \mathbb{R}^4 in Example 2.19, and it is easy to verify that both are $(2, 0)$ tensors (why isn't the Hermitian scalar product of Example 2.21 included?). We also have the inverse metrics, defined in Problem 2.7, and you can verify that these are $(0, 2)$ tensors.

Exercise 3.3 Show that for a metric g on V ,

$$g_i^j = \delta_i^j, \quad (3.13)$$

so the $(1, 1)$ tensor associated to g (via $g!$) is just the identity operator. You will need the components g^{ij} of the inverse metric, defined in Problem 2.7 or Example 2.22.

3.2 Change of Basis

Now we are in a position to derive the usual transformation laws that historically were taken as the *definition* of a tensor. Suppose we have a vector space V and two bases for V , $\mathcal{B} = \{e_i\}_{i=1, \dots, n}$ and $\mathcal{B}' = \{e_{i'}\}_{i'=1, \dots, n}$. Since \mathcal{B} is a basis, each of the $e_{i'}$ can be expressed as

$$e_{i'} = A_{i'}^j e_j \quad (3.14)$$

for some numbers $A_{i'}^j$. Likewise, there exist numbers $A_i^{j'}$ (note that here the upper index is primed) such that

$$e_i = A_i^{j'} e_{j'}. \quad (3.15)$$

We then have

$$e_i = A_i^{j'} e_{j'} = A_i^{j'} A_{j'}^k e_k \quad (3.16)$$

and can then conclude that

$$A_i^{j'} A_{j'}^k = \delta_i^k. \quad (3.17)$$

Considering (3.16) with the primed and unprimed indices switched also yields

$$A_{i'}^j A_j^{k'} = \delta_{i'}^{k'}, \quad (3.18)$$

so, in a way, $A_i^{j'}$ and $A_{i'}^j$ are inverses of each other. Notice that $A_{i'}^j$ and $A_i^{j'}$ are *not* to be interpreted as the components of tensors, as their indices refer to different bases.² How do the corresponding dual bases transform? Let $\{e^i\}_{i=1,\dots,n}$ and $\{e^{i'}\}_{i=1,\dots,n}$ be the bases dual to \mathcal{B} and \mathcal{B}' . Then the components of $e^{i'}$ with respect to $\{e^i\}_{i=1,\dots,n}$ are

$$e^{i'}(e_j) = e^{i'}(A_j^{k'} e_{k'}) = A_j^{k'} \delta_{k'}^{i'} = A_j^{i'}, \quad (3.19)$$

i.e.

$$e^{i'} = A_j^{i'} e^j. \quad (3.20)$$

Likewise,

$$e^i = A_{j'}^i e^{j'}. \quad (3.21)$$

Notice how well the Einstein summation convention and our convention for priming indices work together in the transformation laws. Now we are ready to see how the components of a general (r, s) tensor T transform:

$$\begin{aligned} T_{i'_1, \dots, i'_r}^{j'_1 \dots j'_s} &= T(e_{i'_1}, \dots, e_{i'_r}, e^{j'_1}, \dots, e^{j'_s}) \\ &= T(A_{i'_1}^{k_1} e_{k_1}, \dots, A_{i'_r}^{k_r} e_{k_r}, A_{l'_1}^{j'_1} e^{l_1}, \dots, A_{l'_s}^{j'_s} e^{l_s}) \\ &= A_{i'_1}^{k_1} \dots A_{i'_r}^{k_r} A_{l'_1}^{j'_1} \dots A_{l'_s}^{j'_s} T(e_{k_1}, \dots, e_{k_r}, e^{l_1}, \dots, e^{l_s}) \\ &= \boxed{A_{i'_1}^{k_1} \dots A_{i'_r}^{k_r} A_{l'_1}^{j'_1} \dots A_{l'_s}^{j'_s} T_{k_1 \dots k_r}^{l_1 \dots l_s}}. \end{aligned} \quad (3.22)$$

Equation (3.22) is the standard tensor transformation law, which is taken as the *definition* of a tensor in much of the physics literature; here we have *derived* it as a consequence of our definition of a tensor as a multilinear function on V and V^* . The two are equivalent, however, as you will check in Exercise 3.4 below. With the general transformation law in hand, we will now look at specific types of tensors and

²This is also why we wrote the upper index directly above the lower index, rather than with a horizontal displacement as is customary for tensors. For more about these numbers and a possible interpretation, see the beginning of the next section.

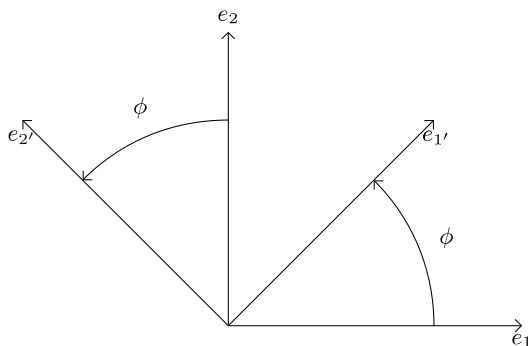


Fig. 3.2 The standard basis \mathcal{B} and a new one \mathcal{B}' obtained by rotation through an angle θ

derive their *matrix* transformation laws; to this end, it will be useful to introduce the matrices

$$A = \begin{pmatrix} A_1^{1'} & A_2^{1'} & \dots & A_n^{1'} \\ A_1^{2'} & A_2^{2'} & \dots & A_n^{2'} \\ \vdots & \vdots & \ddots & \vdots \\ A_1^{n'} & A_2^{n'} & \dots & A_n^{n'} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} A_{1'}^1 & A_{2'}^1 & \dots & A_{n'}^1 \\ A_{1'}^2 & A_{2'}^2 & \dots & A_{n'}^2 \\ \vdots & \vdots & \ddots & \vdots \\ A_{1'}^n & A_{2'}^n & \dots & A_{n'}^n \end{pmatrix}. \quad (3.23)$$

By virtue of (3.17) and (3.18), these matrices satisfy

$$AA^{-1} = A^{-1}A = I \quad (3.24)$$

as our notation suggests.

Exercise 3.4 Consider a function which assigns to a basis $\{e_i\}_{i=1,\dots,n}$ a set of numbers $\{T_{k_1 \dots k_r}^{l_1 \dots l_s}\}$ which transform according to (3.22) under a change of basis. Use this assignment to define a multilinear function T of type (r, s) on V , and be sure to check that your definition is basis-independent (i.e. that the value of T does not depend on which basis $\{e_i\}_{i=1,\dots,n}$ you choose).

Example 3.6 Change of basis matrix for a 2-D rotation

As a simple illustration of the formalism, consider the standard basis \mathcal{B} in \mathbb{R}^2 and another basis \mathcal{B}' obtained by rotating \mathcal{B} by an angle θ . This is illustrated in Fig. 3.2.

By inspection we have

$$e_{1'} = \cos \theta e_1 + \sin \theta e_2 \quad (3.25)$$

$$e_{2'} = -\sin \theta e_1 + \cos \theta e_2 \quad (3.26)$$

and so by (3.14) we have

$$\begin{aligned} A_{1'}^1 &= \cos \theta, & A_{2'}^1 &= -\sin \theta \\ A_{1'}^2 &= \sin \theta, & A_{2'}^2 &= \cos \theta. \end{aligned}$$

Equation (3.23) then tells us that

$$A^{-1} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

The numbers $A_i^{j'}$ and the corresponding matrix A can be computed by either inverting A^{-1} , or equivalently by inverting the system (3.25)–(3.26) and proceeding as above. \square

Example 3.7 Vectors and dual vectors

Given a vector v (considered as a $(0, 1)$ tensor as per (3.3)), (3.22) tells us that its components transform as

$$v^{i'} = A_j^{i'} v^j \quad (3.27)$$

while the components of a dual vector f transform as

$$f_{i'} = A_{i'}^j f_j. \quad (3.28)$$

Notice that the components of v transform with the $A_j^{i'}$ whereas the basis vectors transform with the $A_{i'}^j$, so the components of a vector obey the law opposite ('contra-') to the basis vectors. **This is the origin of the term 'contravariant'**. Note also that the components of a dual vector transform in the *same* way as the basis vectors, **hence the term 'covariant'**. It makes sense that the basis vectors and the components of a vector should transform oppositely; v exists independently of any basis for V and should not change under a change of basis, so if the e_i change one way the v^i should change oppositely. Similar remarks apply to dual vectors.

Aside Incidentally, we can now explain a little bit more about the Einstein summation convention. We knew ahead of time that the components of dual vectors would transform like basis vectors, so we gave them both lower indices. We also knew that the components of vectors would transform like dual basis vectors, so we gave them both upper indices. Since the two transformation laws are opposite, we know (see below) that a summation over an upper index and lower index will yield an object that does not transform at all, so the summation represents an object on a process that is invariant, i.e. the expression $v^i e_i$ represents the vector v , which is defined without reference to any basis. Also, the expression $f_i v^i$ is just $f(v)$, the action of the functional f on the vector v , also defined without reference to any basis. Processes such as these are so important and ubiquitous that it becomes very convenient to omit the summation sign for repeated upper and lower indices, and we thus have the summation convention. Occasionally one encounters two repeated upper indices or two repeated lower indices that are to be summed over; we choose to indicate summation explicitly in these cases, rather than conventionally omit the summation sign, because in these cases there is usually some assumption at work, e.g. the assumption of an orthonormal basis (as in (3.32) below). In such cases the formulas are not completely general, and only represent invariant processes when the accompanying assumption is satisfied.

Returning to our discussion of how components of vectors and dual vectors transform, we can write (3.27) and (3.28) in terms of matrices as

$$[v]_{B'} = A[v]_B \quad (3.29)$$

$$[f]_{B'} = A^{-1T} [f]_B \quad (3.30)$$

where the superscript T again denotes the transpose of a matrix. From the ‘aside’ above, we know that $f(v)$ is basis-independent, but we also know that $f(v) = [f]_{\mathcal{B}}^T[v]_{\mathcal{B}}$. This last equation then must be true in *any* basis, and we can in fact prove this using (3.29) and (3.30): in a new basis \mathcal{B}' , we have

$$\begin{aligned} [f]_{\mathcal{B}'}^T[v]_{\mathcal{B}'} &= (A^{-1T} [f]_{\mathcal{B}})^T A[v]_{\mathcal{B}} \\ &= [f]_{\mathcal{B}}^T A^{-1} A[v]_{\mathcal{B}} \\ &= [f]_{\mathcal{B}}^T[v]_{\mathcal{B}}. \end{aligned} \quad (3.31)$$

This makes concrete our claim above that $[f]$ transforms ‘oppositely’ to $[v]$, so that the basis-independent object $f(v)$ really *is* invariant under a change of basis.

Before moving on to our next example we should point out a minor puzzle: you showed in Exercise 2.18 that if we have an inner product $(\cdot | \cdot)$ on a real vector space V and an orthonormal basis $\{e_i\}_{i=1,\dots,n}$ then the components of vectors and their corresponding dual vectors are identical, which is why we were able to ignore the distinction between them for so long. Equations (3.29) and (3.30) seem to contradict this, however, since it looks like the components of dual vectors transform very differently from the components of vectors. How do we explain this? Well, if we change from one *orthonormal* basis to another, we have

$$\delta_{i'j'} = (e_{i'} | e_{j'}) = A_{i'}^k A_{j'}^l (e_k | e_l) = \sum_{k=1}^n A_{i'}^k A_{j'}^k \quad (3.32)$$

which in matrices reads

$$A^{-1T} A^{-1} = I$$

so we must have

$$A^{-1T} = A$$

$$\iff \boxed{A^{-1} = A^T}.$$

Such matrices are known as **orthogonal** matrices, and we see here that a transformation from one orthonormal basis to another is always implemented by an orthogonal matrix.³ For such matrices (3.29) and (3.30) are identical, resolving our contradiction.

Incidentally, for a *complex* inner product space you will show that orthonormal basis changes are implemented by matrices satisfying $A^{-1} = A^\dagger$. Such matrices are known as **unitary** matrices and should be familiar from quantum mechanics.

Exercise 3.5 Show that for any invertible matrix A , $(A^{-1})^T = (A^T)^{-1}$, justifying the sloppiness of our notation above.

Exercise 3.6 Show that for a *complex* inner product space V , the matrix A implementing an orthonormal change of basis satisfies $A^{-1} = A^\dagger$.

³See Problem 3.1 for more on orthogonal matrices, as well as Chap. 4.

Example 3.8 Linear operators

We already noted that linear operators can be viewed as $(1, 1)$ tensors as per (3.4). (3.22) then tells us that, for a linear operator T on V ,

$$T_{i'j'} = A_{i'}^k A_l^{j'} T_k^l$$

which in matrix form reads

$$[T]_{\mathcal{B}'} = A[T]_{\mathcal{B}}A^{-1} \quad (3.33)$$

which is the familiar *similarity transformation* of matrices. This, incidentally, allows us to extend the trace functional from $n \times n$ matrices to linear operators as follows: Given $T \in \mathcal{L}(V)$ and a basis \mathcal{B} for V , define the trace of T as

$$\text{Tr}(T) \equiv \text{Tr}([T]_{\mathcal{B}}).$$

You can then use (3.33) to show (see Exercise 3.9) that $\text{Tr}(T)$ does not depend on the choice of basis \mathcal{B} .

Exercise 3.7 Show that for $v \in V$, $f \in V^*$, $T \in \mathcal{L}(V)$, $f(Tv) = [f]^T [T][v]$ is invariant under a change of basis. Use the matrix transformation laws as we did in (3.31).

Exercise 3.8 Let $\mathcal{B} = \{x, y, z\}$, $\mathcal{B}' = \{x + iy, z, x - iy\}$ be bases for $\mathcal{H}_1(\mathbb{R}^3)$, and consider the operator L_z for which matrix expressions were found with respect to both bases in Example 2.15. Find the numbers $A_{j'}^{i'}$ and A_l^j and use these, along with (3.33), to obtain $[L_z]_{\mathcal{B}'}$ from $[L_z]_{\mathcal{B}}$.

Exercise 3.9 Show that (3.33) implies that $\text{Tr}([T]_{\mathcal{B}})$ does not depend on the choice of basis \mathcal{B} , so that $\text{Tr}(T)$ is well-defined.

Example 3.9 $(2, 0)$ tensors

$(2, 0)$ tensors g , which include important examples such as the Minkowski metric and the Euclidean metric, transform as follows according to (3.22):

$$g_{i'j'} = A_{i'}^k A_{j'}^l g_{kl}$$

or in matrix form

$$[g]_{\mathcal{B}'} = A^{-1T} [g]_{\mathcal{B}} A^{-1}. \quad (3.34)$$

Notice that if g is an inner product and \mathcal{B} and \mathcal{B}' are orthonormal bases then $[g]_{\mathcal{B}'} = [g]_{\mathcal{B}} = I$ and (3.34) becomes

$$I = A^{-1T} A^{-1},$$

again telling us that A must be orthogonal. Also note that if A is orthogonal, (3.34) is identical to (3.33), so we do not have to distinguish between $(2, 0)$ tensors and linear operators (as most of us have not in the past!). In the case of the Minkowski

metric η we are not dealing with an inner product but we do have orthonormal bases, with respect to which⁴ η takes the form

$$[\eta] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (3.35)$$

If we are changing from one orthonormal basis to another we then have

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = A^{-1T} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} A^{-1} \quad (3.36)$$

or equivalently

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = A^T \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} A. \quad (3.37)$$

Matrices A satisfying (3.37) are known as **Lorentz Transformations**. Notice that these matrices are not quite orthogonal, so the components of vectors will transform slightly differently than those of dual vectors under these transformations. This is in contrast to the case of \mathbb{R}^n with a positive-definite metric, where if we go from one orthonormal basis to another, then the components of vectors and dual vectors transform identically, as you showed in Exercise 2.18. \square

Exercise 3.10 As in previous exercises, show using the matrix transformation laws that $g(v, w) = [w]^T [g][v]$ is invariant under a change of basis.

3.3 Active and Passive Transformations

Before we move on to the tensor product, we have a little unfinished business to conclude. In the last section when we said that the $A_i^{j'}$ were not the components of a tensor, we were lying a little; there *is* a tensor lurking around, namely the linear operator U that takes the new basis vectors into the old, i.e. $U(e_{i'}) = e_i \forall i$ (the action of U on an arbitrary vector is then given by expanding that vector in the basis \mathcal{B}' and using linearity). What are the components of this tensor? Well, in the old basis \mathcal{B} we have

$$\begin{aligned} U_i{}^j &= U(e_i, e^j) = e^j(Ue_i) = e^j(U(A_i^{k'} e_{k'})) = A_i^{k'} e^j(U(e_{k'})) = A_i^{k'} e^j(e_k) \\ &= A_i^{j'} \end{aligned} \quad (3.38)$$

⁴We assume here that the basis vector e_t satisfying $\eta(e_t, e_t) = -1$ is the fourth vector in the basis, which is not necessary but is somewhat conventional in physics.

so the $A_i^{j'}$ actually are the components of a tensor!⁵ Why did we lie, then? Well, the approach we have been taking so far is to try and think about things in a basis-independent way, and although U is a well-defined linear operator, its definition depends entirely on the two bases we have chosen, so we may as well work directly with the numbers that relate the bases. Also, using one primed index and one unprimed index makes it easy to remember transformation laws like (3.20) and (3.21), but is not consistent with our notation for the components of tensors.

If we write out the components of U as a matrix, you should verify that

$$[e_i]_{\mathcal{B}} = [U]_{\mathcal{B}}[e_{i'}]_{\mathcal{B}} = A[e_{i'}]_{\mathcal{B}} \quad (3.39)$$

which should be compared to (3.29), which reads $[v]_{\mathcal{B}'} = A[v]_{\mathcal{B}}$. Equation (3.39) is called an **active** transformation, since we use the matrix A to change one vector into another, namely $e_{i'}$ into e_i . Note that in (3.39) all vectors are expressed in the *same* basis. Equation (3.29), on the other hand, is called a **passive** transformation, since we use the matrix A not to change the vector v but rather to change the basis which v is referred to, hence changing its components. The notation in most physics texts is not as explicit as ours; one usually sees matrix equations like

$$\mathbf{r}' = A\mathbf{r} \quad (3.40)$$

for both passive and active transformations, and one must rely on context to figure out how the equation is to be interpreted. In the active case, one considers the coordinate system fixed and interprets the matrix A as taking the physical vector \mathbf{r} into a *new* vector \mathbf{r}' , where the components of both are expressed in the *same* coordinate system, just as in (3.39). In the passive case, the physical vector \mathbf{r} does not change but the basis does, so one interprets the matrix A as taking the components of \mathbf{r} in the old coordinate system and giving back the components of the *same* vector \mathbf{r} in the new (primed) coordinate system, just as in (3.29). All this is illustrated in Fig. 3.3.

Before we get to some examples, note that in the passive transformation (3.29) the matrix A takes the old components to the new components, whereas in the active transformation (3.39) A takes the *new* basis vectors to the old ones. Thus when A is interpreted actively it corresponds to the *opposite* transformation as in the passive case. This dovetails with the fact that components and basis vectors transform oppositely, as discussed under (3.28).

Example 3.10 Active and passive orthogonal transformations in two dimensions

Let $\mathcal{B} = \{e_1, e_2\}$ be the standard basis for \mathbb{R}^2 , and consider a new basis \mathcal{B}' given by

$$e_{1'} \equiv \frac{1}{\sqrt{2}}e_1 + \frac{1}{\sqrt{2}}e_2$$

⁵If the sleight-of-hand with the primed and unprimed indices in the last couple steps of (3.38) bothers you, puzzle it out and see if you can understand it. It may help to note that the prime on an index does not change its numerical value, it is just a reminder that it refers to the primed basis.

$$e_{2'} \equiv -\frac{1}{\sqrt{2}}e_1 + \frac{1}{\sqrt{2}}e_2.$$

You can show (as in Example 3.6) that this leads to an orthogonal change of basis matrix given by

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (3.41)$$

which corresponds to rotating our basis counterclockwise by $\phi = 45^\circ$, see Fig. 3.3(a).

Now consider the vector $\mathbf{r} = \frac{1}{2\sqrt{2}}e_1 + \frac{1}{2\sqrt{2}}e_2$, also depicted in the figure. In the standard basis we have

$$[\mathbf{r}]_B = \begin{pmatrix} \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \end{pmatrix}.$$

What does \mathbf{r} look like in our new basis? From the figure we see that \mathbf{r} is proportional to $e_{1'}$, and that is indeed what we find; using (3.29) we have

$$[\mathbf{r}]_{B'} = A[\mathbf{r}]_B = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} \quad (3.42)$$

as expected. Remember that the column vector at the end of (3.42) is expressed in the *primed* basis.

This was the passive interpretation of (3.40); what about the active interpretation? Taking our matrix A and interpreting it as a linear operator represented in the standard basis we again have

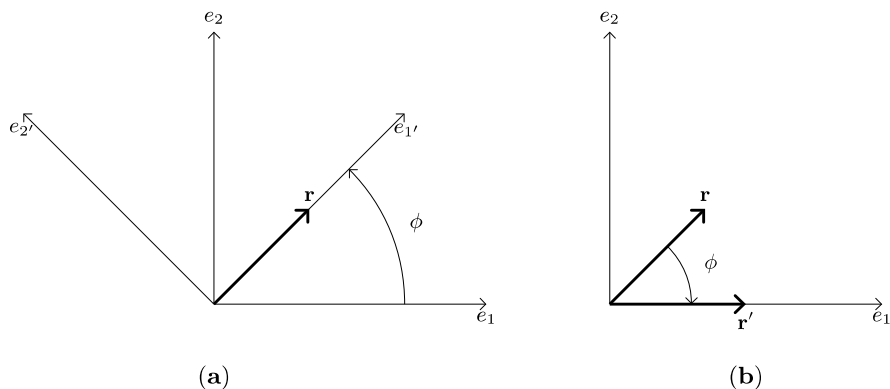


Fig. 3.3 Illustration of the passive and active interpretations of $\mathbf{r}' = A\mathbf{r}$ in two dimensions. In (a) we have a passive transformation, in which the same vector \mathbf{r} is referred to two different bases. The coordinate representation of \mathbf{r} transforms as $\mathbf{r}' = A\mathbf{r}$, though the vector itself does not change. In (b) we have an active transformation, where there is only one basis and the vector \mathbf{r} is itself transformed by $\mathbf{r}' = A\mathbf{r}$. In the active case the transformation is opposite that of the passive case

$$[\mathbf{r}']_{\mathcal{B}} = A[\mathbf{r}]_{\mathcal{B}} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} \quad (3.43)$$

except that now the vector $(1/2, 0)$ represents the *new* vector \mathbf{r}' in the *same* basis \mathcal{B} . This is illustrated in Fig. 3.3(b). As mentioned above, when A is interpreted actively, it corresponds to a *clockwise* rotation, opposite to its interpretation as a passive transformation. \square

Exercise 3.11 Verify (3.41) and (3.42).

Example 3.11 *Active transformations and rigid body motion*

Passive transformations are probably the ones encountered most often in classical physics, since a change of cartesian coordinates induces a passive transformation. Active transformations do crop up, though, especially in the case of rigid body motion. In this scenario, one specifies the orientation of a rigid body by the time-dependent orthogonal basis transformation $A(t)$ which relates the space frame K' to the body frame $K(t)$ (we use here the notation of Example 2.11). As we saw above, there corresponds to the time-dependent matrix $A(t)$ a time-dependent linear operator $U(t)$ which satisfies $U(t)(e_{i'}) = e_i(t)$. If K and K' were coincident at $t = 0$ and \mathbf{r}_0 is the position vector of a point p of the rigid body at that time (see Fig. 3.4(a)), then the position of p at a later time is just $\mathbf{r}(t) = U(t)\mathbf{r}_0$ (see Fig. 3.4(b)), which as a matrix equation in K' would read

$$[\mathbf{r}(t)]_{K'} = A(t)[\mathbf{r}_0]_{K'}. \quad (3.44)$$

In more common and less precise notation this would be written

$$\mathbf{r}(t) = A(t)\mathbf{r}_0.$$

In other words, the position of a specific point on the rigid body at an arbitrary time t is given by the active transformation corresponding to the matrix $A(t)$. \square

Example 3.12 *Active and passive transformations and the Schrödinger and Heisenberg pictures*

The duality between passive and active transformations is also present in quantum mechanics. In the Schrödinger picture, one considers observables like the momentum or position operator as acting on the state ket while the basis kets remain fixed. This is the active viewpoint. In the Heisenberg picture, however, one considers the state ket to be fixed and considers the *observables* to be time-dependent (recall that (2.12) is the equation of motion for these operators). Since the operators are time-dependent, their eigenvectors (which form a basis⁶) are time-dependent as well, so this picture is the passive one in which the vectors do not change but the basis does.

⁶For details on why the eigenvectors of Hermitian operators form a basis, at least in the finite-dimensional case, see Hoffman and Kunze [10].

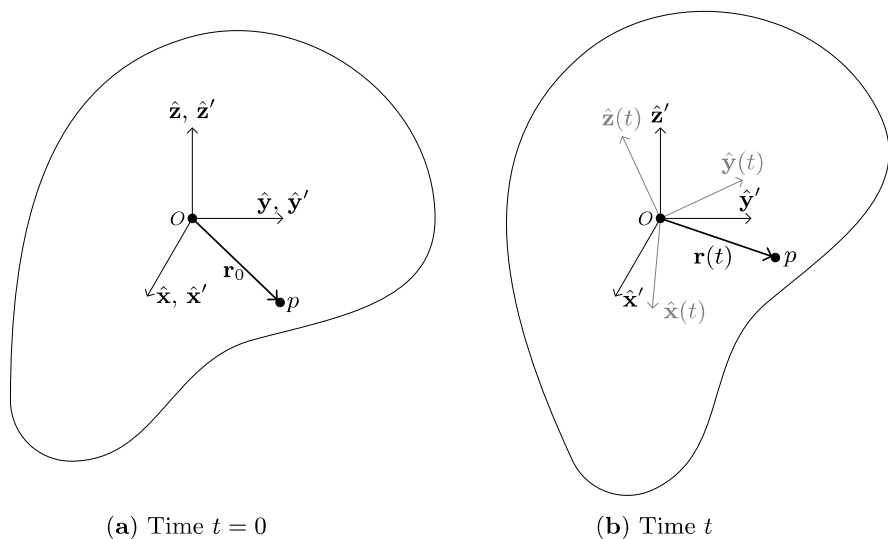


Fig. 3.4 In (a) we have the coincident body and space frames at $t = 0$, along with the point p of the rigid body. In (b) we have the rotated rigid body, and the vector $\mathbf{r}(t)$ pointing to point p now has different components in the space frame, given by (3.44)

Just as an equation like (3.40) can be interpreted in both the active and passive sense, a quantum mechanical equation like

$$\langle \hat{x}(t) \rangle = \langle \psi | (U^\dagger \hat{x} U) | \psi \rangle \quad (3.45)$$

$$= (\langle \psi | U^\dagger) \hat{x} (U | \psi \rangle), \quad (3.46)$$

where U is the time-evolution operator for time t , can also be interpreted in two ways: in the active sense of (3.46), in which the U s act on the vectors and change them into new vectors, and in the passive sense of (3.45), where the U s act on the operator \hat{x} by a similarity transformation to turn it into a new operator, $\hat{x}(t)$.

3.4 The Tensor Product—Definition and Properties

One of the most basic operations with tensors, again commonplace in physics but often unacknowledged (or, at best, dealt with in an ad hoc fashion) is that of the tensor product. Before giving the precise definition, which takes a little getting used to, we give a rough, heuristic description. Given two finite-dimensional vector spaces V and W (over the same set of scalars C), we would like to construct a product vector space, which we denote $V \otimes W$, whose elements are in some sense ‘products’ of vectors $v \in V$ and $w \in W$. We denote these products by $v \otimes w$. This product, like

any respectable product, should be bilinear in the sense that

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \quad (3.47)$$

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2 \quad (3.48)$$

$$c(v \otimes w) = (cv) \otimes w = v \otimes (cw), \quad c \in C. \quad (3.49)$$

Given these properties, the product of any two arbitrary vectors v and w can then be expanded in terms of bases $\{e_i\}_{i=1,\dots,n}$ and $\{f_j\}_{j=1,\dots,m}$ for V and W as

$$\begin{aligned} v \otimes w &= (v^i e_i) \otimes (w^j f_j) \\ &= v^i w^j e_i \otimes f_j \end{aligned}$$

so $\{e_i \otimes f_j\}$, $i = 1, \dots, n$, $j = 1, \dots, m$ should be a basis for $V \otimes W$, which would then have dimension nm . Thus the basis for the product space would be just the product of the basis vectors, and the dimension of the product space would be just the product of the dimensions.

Now we make this precise. Given two finite-dimensional vector spaces V and W , we define their **tensor product** $V \otimes W$ to be the set of all C -valued bilinear functions on $V^* \times W^*$. Such functions do form a vector space, as you can easily check. This definition may seem unexpected or counterintuitive at first, but you will soon see that this definition does yield the vector space described above. Also, given two vectors $v \in V$, $w \in W$, we define their **tensor product** $v \otimes w$ to be the element of $V \otimes W$ defined as follows:

$$(v \otimes w)(h, g) \equiv v(h)w(g) \quad \forall h \in V^*, g \in W^*. \quad (3.50)$$

(Remember that an element of $V \otimes W$ is a bilinear function on $V^* \times W^*$, and so is defined by its action on a pair $(h, g) \in V^* \times W^*$.) The bilinearity of the tensor product is immediate and you can probably verify it without writing anything down: just check that both sides of (3.47)–(3.49) are equal when evaluated on any pair of dual vectors. To prove that $\{e_i \otimes f_j\}$, $i = 1, \dots, n$, $j = 1, \dots, m$ is a basis for $V \otimes W$, let $\{e^i\}_{i=1,\dots,n}$, $\{f^j\}_{j=1,\dots,m}$ be the corresponding dual bases and consider an arbitrary $T \in V \otimes W$. Using bilinearity,

$$T(h, g) = h_i g_j T(e^i, f^j) = h_i g_j T^{ij} \quad (3.51)$$

where $T^{ij} \equiv T(e^i, f^j)$. If we consider the expression $T^{ij} e_i \otimes f_j$, then

$$\begin{aligned} (T^{ij} e_i \otimes f_j)(e^k, f^l) &= T^{ij} e_i(e^k) f_j(f^l) \\ &= T^{ij} \delta_i^k \delta_j^l \\ &= T^{kl} \end{aligned}$$

so $T^{ij} e_i \otimes f_j$ agrees with T on basis vectors, hence on all vectors by bilinearity, so $T = T^{ij} e_i \otimes f_j$. Since T was an arbitrary element of $V \otimes W$, $V \otimes W = \text{Span}\{e_i \otimes f_j\}$. Furthermore, the $e_i \otimes f_j$ are linearly independent as you should check, so $\{e_i \otimes f_j\}$ is actually a basis for $V \otimes W$ and $V \otimes W$ thus has dimension mn .

The tensor product has a couple of important properties besides bilinearity. First, it *commutes with taking duals*, that is,

$$(V \otimes W)^* = V^* \otimes W^*. \quad (3.52)$$

Secondly, and more importantly, the tensor product is *associative*, i.e. for vector spaces V_i , $i = 1, 2, 3$,

$$(V_1 \otimes V_2) \otimes V_3 = V_1 \otimes (V_2 \otimes V_3). \quad (3.53)$$

This property allows us to drop the parentheses and write expressions like $V_1 \otimes \cdots \otimes V_n$ without ambiguity. One can think of $V_1 \otimes \cdots \otimes V_n$ as the set of C -valued multilinear functions on $V_1^* \times \cdots \times V_n^*$.

These two properties are both plausible, particularly when thought of in terms of basis vectors, but verifying them rigorously turns out to be slightly tedious. See Warner [18] for proofs and further details.

Exercise 3.12 If $\{e_i\}$, $\{f_j\}$ and $\{g_k\}$ are bases for V_1 , V_2 and V_3 , respectively, convince yourself that $\{e_i \otimes f_j \otimes g_k\}$ is a basis for $V_1 \otimes V_2 \otimes V_3$, and hence that $\dim V_1 \otimes V_2 \otimes V_3 = n_1 n_2 n_3$ where $\dim V_i = n_i$. Extend the above to n -fold tensor products.

3.5 Tensor Products of V and V^*

In the previous section we defined the tensor product for two arbitrary vector spaces V and W . Often, though, we will be interested in just the iterated tensor product of a vector space and its dual, i.e. in tensor products of the form

$$\underbrace{V^* \otimes \cdots \otimes V^*}_{r \text{ times}} \otimes \underbrace{V \otimes \cdots \otimes V}_{s \text{ times}}. \quad (3.54)$$

This space is of particular interest because it is actually identical to \mathcal{T}_s^r ! In fact, from the previous section we know that the vector space in (3.54) can be interpreted as the set of multilinear functions on

$$\underbrace{V \times \cdots \times V}_{r \text{ times}} \times \underbrace{V^* \times \cdots \times V^*}_{s \text{ times}},$$

but these functions are exactly \mathcal{T}_s^r . Since the space in (3.54) has basis $\mathcal{B}_s^r = \{e^{i_1} \otimes \cdots \otimes e^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}\}$, we can conclude that \mathcal{B}_s^r is a basis for \mathcal{T}_s^r . In fact, we claim that if $T \in \mathcal{T}_s^r$ has components $T_{i_1 \dots i_r}^{j_1 \dots j_s}$, then

$$\boxed{T = T_{i_1 \dots i_r}^{j_1 \dots j_s} e^{i_1} \otimes \cdots \otimes e^{i_r} \otimes e_{j_1} \otimes \cdots \otimes e_{j_s}} \quad (3.55)$$

is the expansion of T in the basis \mathcal{B}_s^r . To prove this, we just need to check that both sides agree when evaluated on an arbitrary set of basis vectors; on the left hand side we get $T(e_{i_1}, \dots, e_{i_r}, e^{j_1}, \dots, e^{j_s}) = T_{i_1, \dots, i_r}^{j_1 \dots j_s}$ by definition, and on the right

hand side we have

$$\begin{aligned}
 & (T_{k_1 \dots k_r}^{l_1 \dots l_s} e^{k_1} \otimes \dots \otimes e^{k_r} \otimes e_{l_1} \otimes \dots \otimes e_{l_s})(e_{i_1}, \dots, e_{i_r}, e^{j_1}, \dots, e^{j_s}) \\
 &= T_{k_1 \dots k_r}^{l_1 \dots l_s} e^{k_1}(e_{i_1}) \dots e^{k_r}(e_{i_r}) e_{l_1}(e^{j_1}) \dots e_{l_s}(e^{j_s}) \\
 &= T_{k_1 \dots k_r}^{l_1 \dots l_s} \delta_{i_1}^{k_1} \dots \delta_{i_r}^{k_r} \delta_{l_1}^{j_1} \dots \delta_{l_s}^{j_s} \\
 &= T_{i_1, \dots, i_r}^{j_1 \dots j_s}
 \end{aligned} \tag{3.56}$$

so our claim is true. Thus, for instance, a $(2, 0)$ tensor like the Minkowski metric can be written as $\eta = \eta_{\mu\nu} e^\mu \otimes e^\nu$. Conversely, a tensor product like $f \otimes g = f_i g_j e^i \otimes e^j \in T_0^2$ thus has components $(f \otimes g)_{ij} = f_i g_j$. Notice that we now have two ways of thinking about components: either as the values of the tensor on sets of basis vectors (as in (3.5)) or as the expansion coefficients in the given basis (as in (3.55)). This duplicity of perspective was pointed out in the case of vectors just above Exercise 2.10, and it is essential that you be comfortable thinking about components in either way.

Exercise 3.13 Compute the dimension of T_s^r .

Exercise 3.14 Let T_1 and T_2 be tensors of type (r_1, s_1) and (r_2, s_2) , respectively, on a vector space V . Show that $T_1 \otimes T_2$ can be viewed as an $(r_1 + r_2, s_1 + s_2)$ tensor, so that the tensor product of two tensors is again a tensor, justifying the nomenclature.

One important operation on tensors which we are now in a position to discuss is that of **contraction**, which is the generalization of the trace functional to tensors of arbitrary rank: Given $T \in T_s^r(V)$ with expansion

$$T = T_{i_1 \dots i_r}^{j_1 \dots j_s} e^{i_1} \otimes \dots \otimes e^{i_r} \otimes e_{j_1} \otimes \dots \otimes e_{j_s} \tag{3.57}$$

we can define a contraction of T to be any $(r-1, s-1)$ tensor resulting from feeding e^i into one of the arguments, e_i into another and then summing over i as implied by the summation convention. For instance, if we feed e_i into the r th slot and e^i into the $(r+s)$ th slot and sum, we get the $(r-1, s-1)$ tensor \tilde{T} defined as

$$\tilde{T}(v_1, \dots, v_{r-1}, f_1, \dots, f_{s-1}) \equiv T(v_1, \dots, v_{r-1}, e_i, f_1, \dots, f_{s-1}, e^i).$$

You may be suspicious that \tilde{T} depends on our choice of basis, but Exercise 3.15 shows that contraction is in fact well-defined. Notice that the components of \tilde{T} are

$$\tilde{T}_{i_1 \dots i_{r-1}}^{j_1 \dots j_{s-1}} = T_{i_1 \dots i_{r-1} l}^{j_1 \dots j_{s-1} l}.$$

Similar contractions can be performed on any two arguments of T provided one argument eats vectors and the other dual vectors. In terms of components, a contraction can be taken with respect to any pair of indices provided that one is covariant and the other contravariant. If we are working on a vector space equipped with a metric g , then we can use the metric to raise and lower indices and so can contract on any pair of indices, even if they are both covariant or contravariant. For instance, we can contract a $(2, 0)$ tensor T with components T_{ij} as $\tilde{T} = T_i^i = g^{ij} T_{ij}$, which one can interpret as just the trace of the associated linear operator (or $(1, 1)$ tensor).

For a linear operator or any other rank 2 tensor, this is the only option for contraction. If we have two linear operators A and B , then their tensor product $A \otimes B \in \mathcal{T}_2^2$ has components

$$(A \otimes B)_{ik}{}^{jl} = A_i{}^j B_k{}^l,$$

and contracting on the first and last index gives a $(1, 1)$ tensor AB whose components are

$$(AB)_k{}^j = A_l{}^j B_k{}^l.$$

You should check that this tensor is just the composition of A and B , as our notation suggests. What linear operator do we get if we consider the other contraction $A_i{}^j B_j{}^l$?

Exercise 3.15 Show that if $\{e_i\}_{i=1,\dots,n}$ and $\{e_{i'}\}_{i=1,\dots,n}$ are two arbitrary bases that

$$T(v_1, \dots, v_{r-1}, e_i, f_1, \dots, f_{s-1}, e^i) = T(v_1, \dots, v_{r-1}, e_{i'}, f_1, \dots, f_{s-1}, e^{i'})$$

so that contraction is well-defined.

Example 3.13 $V^* \otimes V$

One of the most important examples of tensor products of the form (3.54) is $V^* \otimes V$, which as we mentioned is the same as \mathcal{T}_1^1 , the space of linear operators. How does this identification work, explicitly? Well, given $f \otimes v \in V^* \otimes V$, we can define a linear operator by $(f \otimes v)(w) \equiv f(w)v$. More generally, given

$$T_i{}^j e^i \otimes e_j \in V^* \otimes V, \quad (3.58)$$

we can define a linear operator T by

$$T(v) = T_i{}^j e^i(v) e_j = v^i T_i{}^j e_j$$

which is identical to (2.13). This identification of $V^* \otimes V$ and linear operators is actually implicit in many quantum mechanical expressions. Let \mathcal{H} be a quantum mechanical Hilbert space and let $\psi, \phi \in \mathcal{H}$ so that $L(\phi) \in \mathcal{H}^*$. The tensor product of $L(\phi)$ and ψ , which we would write as $L(\phi) \otimes \psi$, is written in Dirac notation as $|\psi\rangle\langle\phi|$. If we are given an orthonormal basis $\mathcal{B} = \{|i\rangle\}$, the expansion (3.58) of an arbitrary operator H can be written in Dirac notation as

$$H = \sum_{i,j} H_{ij} |i\rangle\langle j|,$$

an expression which may be familiar from advanced quantum mechanics texts.⁷ In particular, the identity operator can be written as

$$I = \sum_i |i\rangle\langle i|,$$

which is referred to as the **resolution of the identity** with respect to the basis $\{|i\rangle\}$.

⁷We do not bother here with index positions since most quantum mechanics texts do not employ Einstein summation convention, preferring instead to explicitly indicate summation.

A word about nomenclature: In quantum mechanics and other contexts the tensor product is often referred to as the **direct** or **outer** product. This last term is meant to distinguish it from the inner product, since both the outer and inner products eat a dual vector and a vector (strictly speaking the inner product eats 2 vectors, but remember that with an inner product we may identify vectors and dual vectors) but the outer product yields a linear operator whereas the inner product yields a scalar.

Exercise 3.16 Interpret $e^i \otimes e_j$ as a linear operator, and convince yourself that its matrix representation is

$$[e^i \otimes e_j] = E_{ji}.$$

Recall that E_{ji} is one of the elementary basis matrices introduced way back in Example 2.8, and has a 1 in the j th row and i th column and zeros everywhere else.

3.6 Applications of the Tensor Product in Classical Physics

Example 3.14 Moment of inertia tensor revisited

We took an abstract look at the moment of inertia tensor in Example 3.3; now, armed with the tensor product, we can examine the moment of inertia tensor more concretely. Consider a rigid body with a fixed point O , so that it has only rotational degrees of freedom (O need not necessarily be the center of mass). Let O be the origin, pick time-dependent body-fixed axes $K = \{\hat{\mathbf{x}}(t), \hat{\mathbf{y}}(t), \hat{\mathbf{z}}(t)\}$ for \mathbb{R}^3 , and let g denote the Euclidean metric on \mathbb{R}^3 . Recall that g allows us to define a map L from \mathbb{R}^3 to \mathbb{R}^{3*} . Also, let the i th particle in the rigid body have mass m_i and position vector \mathbf{r}_i with $[\mathbf{r}_i]_K = (x_i, y_i, z_i)$ relative to O , and let $r_i^2 \equiv g(\mathbf{r}_i, \mathbf{r}_i)$. All this is illustrated in Fig. 3.5. The $(2, 0)$ moment of inertia tensor is then given by

$$\mathcal{I}_{(2,0)} = \sum_i m_i (r_i^2 g - L(\mathbf{r}_i) \otimes L(\mathbf{r}_i)) \quad (3.59)$$

while the $(1, 1)$ tensor reads

$$\mathcal{I}_{(1,1)} = \sum_i m_i (r_i^2 I - L(\mathbf{r}_i) \otimes \mathbf{r}_i). \quad (3.60)$$

You should check that in components (3.59) reads

$$\mathcal{I}_{jk} = \sum_i m_i (r_i^2 \delta_{jk} - (\mathbf{r}_i)_j (\mathbf{r}_i)_k).$$

Writing a couple of components explicitly yields

$$\begin{aligned} \mathcal{I}_{xx} &= \sum_i m_i (y_i^2 + z_i^2) \\ \mathcal{I}_{xy} &= - \sum_i m_i x_i y_i, \end{aligned} \quad (3.61)$$

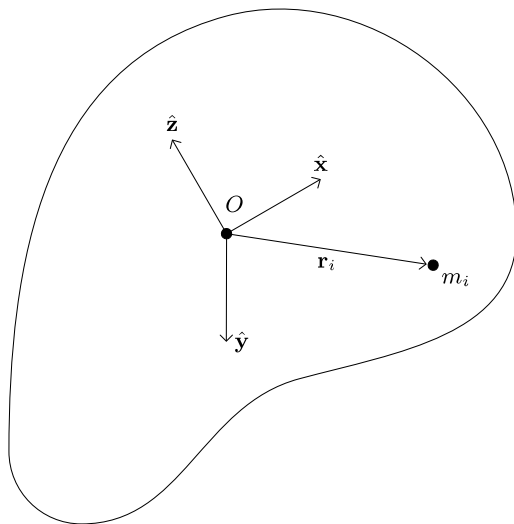


Fig. 3.5 The rigid body with fixed point O and body-fixed axes $K = \{\hat{x}, \hat{y}, \hat{z}\}$, along with i th particle at position \mathbf{r}_i with mass m_i

expressions which should be familiar from classical mechanics. So long as the basis is orthonormal, the components \mathcal{I}_j^k of the $(1, 1)$ tensor in (3.60) will be the same as for the $(2, 0)$ tensor, as remarked earlier. Note that if we had not used body-fixed axes, the components of \mathbf{r}_i (and hence the components of \mathcal{I} , by (3.61)) would in general be time-dependent; this is the main reason for using the body-fixed axes in computation.

Example 3.15 Maxwell stress tensor

In considering the conservation of total momentum (mechanical plus electromagnetic) in classical electrodynamics one encounters the symmetric rank 2 **Maxwell Stress Tensor**, defined in $(2, 0)$ form as⁸

$$T_{(2,0)} = \mathbf{E} \otimes \mathbf{E} + \mathbf{B} \otimes \mathbf{B} - \frac{1}{2}(\mathbf{E} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{B})\mathbf{g}$$

where \mathbf{E} and \mathbf{B} are the dual vector versions of the electric and magnetic field vectors. T can be interpreted in the following way: $T(v, w)$ gives the rate at which momentum in the v -direction flows in the w -direction. In components we have

$$T_{ij} = E_i E_j + B_i B_j - \frac{1}{2}(\mathbf{E} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{B})\delta_{ij},$$

which is the expression found in most classical electrodynamics textbooks.

⁸Recall that we have set all physical constants such as c and ϵ_0 equal to 1.

Example 3.16 *The electromagnetic field tensor*

As you have probably seen in discussions of relativistic electrodynamics, the electric and magnetic field vectors are properly viewed as components of a rank 2 antisymmetric tensor F , the electromagnetic *field tensor*.⁹ To write F in component-free notation requires machinery outside the scope of this text,¹⁰ so we settle for its expression as a matrix in an orthonormal basis, which in $(2, 0)$ form is

$$[F_{(2,0)}] = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \quad (3.62)$$

The Lorentz force law

$$\frac{dp^\mu}{d\tau} = q F^\mu{}_\nu v^\nu$$

where $p = mv$ is the 4-momentum of a particle, v is its 4-velocity and q its charge, can be rewritten without components as

$$\frac{dp}{d\tau} = q F_{(1,1)}(v) \quad (3.63)$$

which just says that the Minkowski force $\frac{dp}{d\tau}$ on a particle is given by the action of the field tensor on the particle's 4-velocity!

3.7 Applications of the Tensor Product in Quantum Physics

In this section we will discuss further applications of the tensor product in quantum mechanics, in particular the oft-unwritten rule that to add degrees of freedom one should take the tensor product of the corresponding Hilbert spaces. Before we get to this, however, we must set up a little more machinery and address an issue that we have so far swept under the rug. The issue is that when dealing with spatial degrees of freedom, as opposed to ‘internal’ degrees of freedom like spin, we often encounter Hilbert spaces like $L^2([-a, a])$ and $L^2(\mathbb{R})$ which are most conveniently described by ‘basis’ vectors which are eigenvectors of either the position operator \hat{x} or the momentum operator \hat{p} . The trouble with these bases is that they are often non-denumerably infinite (i.e. cannot be indexed by the integers, unlike all the bases we have worked with so far) and, what is worse, the ‘basis vectors’ do not even belong

⁹In this example and the one above we are actually not dealing with tensors but with *tensor fields*, i.e. tensor-valued functions on space and spacetime. For the discussion here, however, we will ignore the spatial dependence of the fields, focusing instead on the tensorial properties.

¹⁰One needs the exterior derivative, a generalization of the curl, divergence and gradient operators from vector calculus. See Schutz [15] for a very readable account.

to the Hilbert space! Consider, for example, $L^2(\mathbb{R})$. The position operator \hat{x} acts on functions $\psi(x) \in L^2(\mathbb{R})$ by

$$\hat{x}\psi(x) = x\psi(x). \quad (3.64)$$

If we follow the practice of most quantum mechanics texts and treat the Dirac delta functional δ as $L(\delta(x))$ where $\delta(x)$, the ‘Dirac delta function’, is infinite at 0 and 0 elsewhere, you can check (see Exercise 3.17) that

$$\hat{x}\delta(x - x_0) = x_0\delta(x - x_0)$$

so that $\delta(x - x_0)$ is an ‘eigenfunction’ of \hat{x} with eigenvalue x_0 (in Dirac notation we write the corresponding ket as $|x_0\rangle$). The trouble is that, as we saw in Example 2.23, there is no such $\delta(x) \in L^2(\mathbb{R})$! Furthermore, since the basis $\{\delta(x - x_0)\}_{x_0 \in \mathbb{R}}$ is indexed by \mathbb{R} and not some subset of \mathbb{Z} , we must expand $\psi \in L^2(\mathbb{R})$ by integrating instead of summing. Integration, however, is a limiting procedure and one should really worry about what it means for an integral to converge. Rectifying all this in a rigorous manner is possible,¹¹ but outside the scope of this text, unfortunately. We do wish to work with these objects, however, so we will content ourselves with the traditional approach: ignore the fact that the delta functions are not elements of $L^2(\mathbb{R})$, work without discomfort with the basis $\{\delta(x - x_0)\}_{x_0 \in \mathbb{R}}$,¹² and fearlessly expand arbitrary functions ψ in the basis $\{\delta(x - x_0)\}_{x_0 \in \mathbb{R}}$ as

$$\psi(x) = \int_{-\infty}^{\infty} dx' \psi(x') \delta(x - x'), \quad (3.65)$$

where the above equation can be interpreted both as the expansion of ψ and just the definition of the delta function. In Dirac notation (3.65) reads

$$|\psi\rangle = \int_{-\infty}^{\infty} dx' \psi(x') |x'\rangle. \quad (3.66)$$

Note that we can think of the numbers $\psi(x)$ as the components of $|\psi\rangle$ with respect to the basis $\{|x\rangle\}_{x \in \mathbb{R}}$. Alternatively, if we define the inner product of our basis vectors to be

$$\langle x | x' \rangle \equiv \delta(x - x')$$

as is usually done, then using (3.66) we have

$$\psi(x) = \langle x | \psi \rangle \quad (3.67)$$

which gives another interpretation of $\psi(x)$. These two interpretations of $\psi(x)$ are just the ones mentioned below (3.56); that is, the components of a vector can be interpreted either as expansion coefficients, as in (3.66), or as the value of a given dual vector on the vector, as in (3.67).

¹¹This requires the so-called ‘rigged’ Hilbert space; see Ballentine [2].

¹²Working with the momentum eigenfunctions e^{ipx} instead does not help; though these are legitimate functions, they still are not square-integrable since $\int_{-\infty}^{\infty} |e^{ipx}|^2 dx = \infty$!

Exercise 3.17 By considering the integral

$$\int_{-\infty}^{\infty} (\hat{x}\delta(x-x_0))f(x)dx$$

(where f is an arbitrary square-integrable function), show formally that

$$\hat{x}\delta(x-x_0) = x_0\delta(x-x_0).$$

Exercise 3.18 Check that $\{\delta(x-x_0)\}_{x_0 \in \mathbb{R}}$ satisfies (2.28).

Exercise 3.19 Verify (3.67).

While we are at it, let us pose the following question: we mentioned in a footnote on the previous page that one could use momentum eigenfunctions instead of position eigenfunctions as a basis for $L^2(\mathbb{R})$; what does the corresponding change of basis look like?

Example 3.17 The momentum representation

As is well-known from quantum mechanics, the eigenfunctions of the momentum operator $\hat{p} = -i\frac{d}{dx}$ are the wavefunctions $\{e^{ipx}\}_{p \in \mathbb{R}}$, and these wavefunctions form a basis for $L^2(\mathbb{R})$. In fact, the expansion of an arbitrary function $\psi \in L^2(\mathbb{R})$ in this basis is just the Fourier expansion of ψ , written

$$\psi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \phi(p) e^{ipx} \quad (3.68)$$

where the component function $\phi(p)$ is known as the **Fourier Transform** of ψ . One could in fact work exclusively with $\phi(p)$ instead of $\psi(x)$, and recast the operators \hat{x} and \hat{p} in terms of their action on $\phi(p)$ (see Exercise 3.20 below); such an approach is known as the **momentum representation**. Now, what does it look like when we switch from the position representation to the momentum representation, i.e. when we change bases from $\{\delta(x-x_0)\}_{x_0 \in \mathbb{R}}$ to $\{e^{ipx}\}_{p \in \mathbb{R}}$? Since the basis vectors are indexed by real numbers p and x_0 as opposed to integers i and j , our change of basis will not be given by a matrix with components A_j^i , but rather a *function* $A(x_0, p)$. By (3.19) and the fact that both bases are orthonormal, this function is given by the inner product of $\delta(x-x_0)$ and e^{ipx} . In Dirac notation this would be written as $\langle x_0|p\rangle$, and we have

$$\langle x_0|p\rangle = \int_{-\infty}^{\infty} dx \delta(x-x_0) e^{ipx} = e^{ipx_0},$$

a familiar equation. □

Exercise 3.20 Use (3.68) to show that in the momentum representation, $\hat{p}\phi(p) = p\phi(p)$ and $\hat{x}\phi(p) = i\frac{d\phi}{dp}$.

The next issue to address is that of linear operators: having constructed a new Hilbert space¹³ $\mathcal{H}_1 \otimes \mathcal{H}_2$ out of two Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , can we construct

¹³You may have noticed that we defined tensor products only for finite-dimensional spaces. The definition can be extended to cover infinite-dimensional Hilbert spaces, but the extra technicalities

linear operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$ out of the linear operators on \mathcal{H}_1 and \mathcal{H}_2 ? Well, given linear operators A_i on \mathcal{H}_i , $i = 1, 2$, we can define a linear operator $A_1 \otimes A_2$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$(A_1 \otimes A_2)(v \otimes w) \equiv (A_1 v) \otimes (A_2 w). \quad (3.69)$$

You can check that with this definition, $(A \otimes B)(C \otimes D) = AC \otimes BD$. In most quantum mechanical applications either A_1 or A_2 is the identity, i.e. one considers operators of the form $A_1 \otimes I$ or $I \otimes A_2$. These are often abbreviated as A_1 and A_2 even though they are acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$. We should also mention here that the inner product $(\cdot | \cdot)_{\otimes}$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is just the product of the inner products on $(\cdot | \cdot)_i$ on the \mathcal{H}_i , that is,

$$(v_1 \otimes v_2 | w_1 \otimes w_2)_{\otimes} \equiv (v_1, w_1)_1 \cdot (v_2, w_2)_2.$$

The last subject we should touch upon is that of **vector operators**, which are defined to be sets of operators that transform as three-dimensional vectors under the adjoint action of the total angular momentum operators J_i . That is, a vector operator is a set of operators $\{B_i\}_{i=1-3}$ (often written collectively as \mathbf{B}) that satisfies

$$\text{ad}_{J_i}(B_j) = [J_i, B_j] = i \sum_{k=1}^3 \epsilon_{ijk} B_k \quad (3.70)$$

where ϵ_{ijk} is the familiar Levi-Civita symbol. The three-dimensional position operator $\hat{\mathbf{r}} = \{\hat{x}, \hat{y}, \hat{z}\}$, momentum operator $\hat{\mathbf{p}} = \{\hat{p}_x, \hat{p}_y, \hat{p}_z\}$, and orbital angular momentum operator $\mathbf{L} = \{L_x, L_y, L_z\}$ are all vector operators, as you can check.

Exercise 3.21 For spinless particles, $\mathbf{J} = \mathbf{L} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}$. Expressions for the components may be obtained by expanding the cross product or referencing Example 2.15 and Exercise 2.9. Use these expressions and the canonical commutation relations $[x_i, p_j] = -i\delta_{ij}$ to show that $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}$ and \mathbf{L} are all vector operators.

Now we are finally ready to consider some examples, in which we will take as an axiom that **adding degrees of freedom is implemented by taking tensor products of the corresponding Hilbert spaces**. You will see that this process reproduces familiar results.

Example 3.18 Addition of translational degrees of freedom

Consider a spinless particle constrained to move in one dimension; the quantum mechanical Hilbert space for this system is $L^2(\mathbb{R})$ with basis $\{|x\rangle\}_{x \in \mathbb{R}}$. If we consider a second dimension, call it the y dimension, then this degree of freedom has its own Hilbert space $L^2(\mathbb{R})$ with basis $\{|y\rangle\}_{y \in \mathbb{R}}$. If we allow the particle both degrees of freedom then the Hilbert space for the system is $L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$, with basis $\{|x\rangle \otimes |y\rangle\}_{x, y \in \mathbb{R}}$. An arbitrary ket $|\psi\rangle \in L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$ has expansion

needed do not add any insight to what we are trying to do here, so we omit them. The theory of infinite-dimensional Hilbert spaces falls under the rubric of *functional analysis*, and details can be found, for example, in Reed and Simon [12].

$$|\psi\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \psi(x, y) |x\rangle \otimes |y\rangle$$

with expansion coefficients $\psi(x, y)$. If we iterate this logic, we get in 3 dimensions

$$|\psi\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy dz \psi(x, y, z) |x\rangle \otimes |y\rangle \otimes |z\rangle.$$

If we rewrite $\psi(x, y, z)$ as $\psi(\mathbf{r})$ and $|x\rangle \otimes |y\rangle \otimes |z\rangle$ as $|\mathbf{r}\rangle$ where $\mathbf{r} = (x, y, z)$, then we have

$$|\psi\rangle = \int d^3r \psi(\mathbf{r}) |\mathbf{r}\rangle$$

which is the familiar expansion of a ket in terms of three-dimensional position eigenkets. Such a ket is an element of $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$, which is also denoted as $L^2(\mathbb{R}^3)$.¹⁴

Example 3.19 Two-particle systems

Now consider two spinless particles in three-dimensional space, possibly interacting through some sort of potential. The two-body problem with a $1/r$ potential is a classic example of this. The Hilbert space for such a system is then $L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$, with basis $\{|\mathbf{r}_1\rangle \otimes |\mathbf{r}_2\rangle\}_{\mathbf{r}_i \in \mathbb{R}^3}$. In many textbooks the tensor product symbol is omitted and such basis vectors are written as $|\mathbf{r}_1\rangle |\mathbf{r}_2\rangle$ or even $|\mathbf{r}_1, \mathbf{r}_2\rangle$. A ket $|\psi\rangle$ in this Hilbert space then has expansion

$$|\psi\rangle = \int d^3r_1 \int d^3r_2 \psi(\mathbf{r}_1, \mathbf{r}_2) |\mathbf{r}_1, \mathbf{r}_2\rangle$$

which is the familiar expansion of a ket in a two-particle Hilbert space. One can interpret $\psi(\mathbf{r}_1, \mathbf{r}_2)$ as the probability amplitude of finding particle 1 in position \mathbf{r}_1 and particle 2 in position \mathbf{r}_2 simultaneously.

Example 3.20 Addition of orbital and spin angular momentum

Now consider a spin s particle in three dimensions. As remarked in Example 2.2, the ket space corresponding to the spin degree of freedom is \mathbb{C}^{2s+1} , and one usually takes a basis $\{|m\rangle\}_{-s \leq m \leq s}$ of S_z eigenvectors with eigenvalue m . The total Hilbert space for this system is $L^2(\mathbb{R}^3) \otimes \mathbb{C}^{2s+1}$, and we can take as a basis $\{|\mathbf{r}\rangle \otimes |m\rangle\}$ where $\mathbf{r} \in \mathbb{R}^3$ and $-s \leq m \leq s$. Again, the basis vectors are often written as $|\mathbf{r}\rangle |m\rangle$ or even $|\mathbf{r}, m\rangle$. An arbitrary ket $|\psi\rangle$ then has expansion

¹⁴ $L^2(\mathbb{R}^3)$ is actually defined to be the set of all square-integrable functions on \mathbb{R}^3 , i.e. functions f satisfying

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy dz |f|^2 < \infty.$$

Not too surprisingly, this space turns out to be identical to $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \otimes L^2(\mathbb{R})$.

$$|\psi\rangle = \sum_{m=-s}^s \int d^3r \psi_m(\mathbf{r}) |\mathbf{r}, m\rangle$$

where $\psi_m(\mathbf{r})$ is the probability of finding the particle at position \mathbf{r} and with m units of spin angular momentum in the z -direction. These wavefunctions are sometimes written in column vector form

$$\begin{pmatrix} \psi_s \\ \psi_{s-1} \\ \vdots \\ \psi_{-s+1} \\ \psi_{-s} \end{pmatrix}.$$

The total angular momentum operator \mathbf{J} is given by $\mathbf{L} \otimes I + I \otimes \mathbf{S}$ where \mathbf{L} is the orbital angular momentum operator. One might wonder why \mathbf{J} is not given by $\mathbf{L} \otimes \mathbf{S}$; there is a good answer to this question, but it requires delving into the (fascinating) subject of Lie groups and Lie algebras, which we postpone until part II. In the meantime, you can get a partial answer by checking (Exercise 3.22 below) that the operators $L_i \otimes S_i$ do not satisfy the angular momentum commutation relations whereas the $L_i \otimes I + I \otimes S_i$ do.

Exercise 3.22 Check that

$$[L_i \otimes I + I \otimes S_i, L_j \otimes I + I \otimes S_j] = \sum_{k=1}^3 \epsilon_{ijk} (L_k \otimes I + I \otimes S_k).$$

Also show that

$$[L_i \otimes S_i, L_j \otimes S_j] \neq \sum_{k=1}^3 \epsilon_{ijk} L_k \otimes S_k.$$

Be sure to use the bilinearity of the tensor product carefully.

Example 3.21 Addition of spin angular momentum

Next consider two particles of spin s_1 and s_2 , respectively, fixed in space so that they have no translational degrees of freedom. The Hilbert space for this system is $\mathbb{C}^{2s_1+1} \otimes \mathbb{C}^{2s_2+1}$, with basis $\{|m_1\rangle \otimes |m_2\rangle\}$ where $-s_i \leq m_i \leq s_i$, $i = 1, 2$. Again, such tensor product kets are usually abbreviated as $|m_1\rangle|m_2\rangle$ or $|m_1, m_2\rangle$. There are several important linear operators on $\mathbb{C}^{2s_1+1} \otimes \mathbb{C}^{2s_2+1}$:

- $\mathbf{S}_1 \otimes I$ Vector spin operator on first particle
- $I \otimes \mathbf{S}_2$ Vector spin operator on second particle
- $\mathbf{S} \equiv \mathbf{S}_1 \otimes I + I \otimes \mathbf{S}_2$ Total vector spin operator
- $\mathbf{S}^2 \equiv \sum_i S_i S_i$ Total spin squared operator

(Why aren't \mathbf{S}_1^2 and \mathbf{S}_2^2 in our list above?) The vectors $|m_1, m_2\rangle$ are clearly eigenvectors of S_{1z} and S_{2z} and hence S_z (we abuse notation as mentioned below (3.70)) but, as you will show in Exercise 3.23, they are *not* necessarily eigenvectors of \mathbf{S}^2 .

However, since the S_i obey the angular momentum commutation relations (as you can check), the general theory of angular momentum tells us that we *can* find a basis for $\mathbb{C}^{2s_1+1} \otimes \mathbb{C}^{2s_2+1}$ consisting of eigenvectors of S_z and \mathbf{S}^2 . Furthermore, it can be shown that the \mathbf{S}^2 eigenvalues that occur are $s(s+1)$ where

$$s = |s_1 - s_2|, |s_1 - s_2| + 1, \dots, s_1 + s_2 \quad (3.71)$$

and for a given s the possible S_z eigenvalues are m where $-s \leq m \leq s$ as usual (see Sakurai [14] for details). We will write these basis kets as $\{|s, m\rangle\}$ where the above restrictions on s and m are understood, and where we physically interpret $\{|s, m\rangle\}$ as a state with total angular momentum equal to $\sqrt{s(s+1)}$ and with m units of angular momentum pointing along the z -axis. Thus we have two natural and useful bases for $\mathbb{C}^{2s_1+1} \otimes \mathbb{C}^{2s_2+1}$:

$$\begin{aligned} \mathcal{B} &= \{|m_1, m_2\rangle\}, \quad -s_1 \leq m_1 \leq s_1, \quad -s_2 \leq m_2 \leq s_2 \\ \mathcal{B}' &= \{|s, m\rangle\}, \quad |s_1 - s_2| \leq s \leq s_1 + s_2, \quad -s \leq m \leq s. \end{aligned}$$

What does the transformation between these two bases look like? Well, by their definition, the $A_{i'}^j$ relating the two bases are given by $e^j(e_{i'})$; using s, m collectively in lieu of the primed index and m_1, m_2 collectively in lieu of the unprimed index, we have, in Dirac notation,

$$A_{s,m}^{m_1,m_2} = \langle m_1, m_2 | s, m \rangle. \quad (3.72)$$

These numbers, the notation for which varies widely throughout the literature, are known as **Clebsch–Gordan Coefficients**. Methods for computing them can be found in any standard quantum mechanics textbook.

Let us illustrate the foregoing with an example. Take two spin 1 particles, so that $s_1 = s_2 = 1$. The Hilbert space for the first particle is \mathbb{C}^3 , with S_{1z} eigenvector basis $\{|-1\rangle, |0\rangle, |1\rangle\}$, and so the two-particle system has nine-dimensional Hilbert space $\mathbb{C}^3 \otimes \mathbb{C}^3$ with corresponding basis

$$\begin{aligned} \mathcal{B} &= \{|i\rangle|j\rangle \mid i, j = -1, 0, 1\} \\ &= \{|1\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|-1\rangle, |0\rangle|1\rangle, \text{etc.}\}. \end{aligned}$$

There should also be another basis consisting of S_z and \mathbf{S}^2 eigenvectors, however. From (3.71) we know that the possible s values are $s = 0, 1, 2$, and it is a standard exercise in angular momentum theory to show that the nine (normalized) S_z and \mathbf{S}^2 eigenvectors are

$$\begin{aligned} |s=2, s_z=2\rangle &= |1\rangle|1\rangle \\ |2, 1\rangle &= \frac{1}{\sqrt{2}}(|1\rangle|0\rangle + |0\rangle|1\rangle) \\ |2, 0\rangle &= \frac{1}{\sqrt{6}}(|1\rangle|-1\rangle + 2|0\rangle|0\rangle + |-1\rangle|1\rangle) \\ |2, -1\rangle &= \frac{1}{\sqrt{2}}(|-1\rangle|0\rangle + |0\rangle|-1\rangle) \end{aligned}$$

$$\begin{aligned}
|2, -2\rangle &= |-1\rangle|-1\rangle \\
|1, 1\rangle &= \frac{1}{\sqrt{2}}(|1\rangle|0\rangle - |0\rangle|1\rangle) \\
|1, 0\rangle &= \frac{1}{\sqrt{2}}(|1\rangle|-1\rangle - |-1\rangle|1\rangle) \\
|1, -1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle|-1\rangle - |-1\rangle|0\rangle) \\
|0, 0\rangle &= \frac{1}{\sqrt{3}}(|0\rangle|0\rangle - |1\rangle|-1\rangle - |-1\rangle|1\rangle).
\end{aligned}$$

These vectors can be found using standard techniques from the theory of ‘addition of angular momentum’; for details see Gasiorowicz [5] or Sakurai [14]. The coefficients appearing on the right hand side of the above equations are precisely the Clebsch–Gordan coefficients (3.72), as a moment’s thought should show.

Exercise 3.23 Show that

$$\mathbf{S}^2 = \mathbf{S}_1^2 \otimes I + I \otimes \mathbf{S}_2^2 + 2 \sum_i S_{1i} \otimes S_{2i}.$$

The right hand side of the above equation is usually abbreviated as $\mathbf{S}_1^2 + \mathbf{S}_2^2 + 2\mathbf{S}_1 \cdot \mathbf{S}_2$. Use this to show that $|m_1, m_2\rangle$ is *not* generally an eigenvector of \mathbf{S}^2 .

Example 3.22 Entanglement

Consider two Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 and their tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. Only some of the vectors in $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as $\psi \otimes \phi$; such vectors are referred to as *separable states* or *product states*. All other vectors must be written as linear combinations of the form $\sum_i \psi_i \otimes \phi_i$, and these vectors are said to be *entangled*, since in this case the measurement of the degrees of freedom represented by \mathcal{H}_1 will influence the measurement of the degrees of freedom represented by \mathcal{H}_2 . The classic example of an entangled state comes from the previous example of two fixed particles with spin; taking $s_1 = s_2 = 1/2$ and writing the standard basis for \mathbb{C}^2 as $\{|+\rangle, |-\rangle\}$, we consider the particular state

$$|+\rangle|-\rangle - |-\rangle|+\rangle. \quad (3.73)$$

If an observer measures the first particle to be spin up then a measurement of the second particle’s spin is guaranteed to be spin-down, and vice versa, so measuring one part of the system affects what one will measure for the other part. This is the sense in which the system is entangled. For a product state $\psi \otimes \phi$, there is no such entanglement: a particular measurement of the first particle cannot affect what one measures for the second, since the second particle’s state will be ϕ no matter what. You will check below that (3.73) is not a product state.

Exercise 3.24 Prove that (3.73) cannot be written as $\psi \otimes \phi$ for any $\psi, \phi \in \mathbb{C}^2$. Do this by expanding ψ and ϕ in the given basis and showing that no choice of expansion coefficients for ψ and ϕ will yield (3.73).

3.8 Symmetric Tensors

Given a vector space V there are certain subspaces of $\mathcal{T}_0^r(V)$ and $\mathcal{T}_r^0(V)$ which are of particular interest: the symmetric and antisymmetric tensors. We will discuss symmetric tensors in this section and antisymmetric tensors in the next. A **symmetric** $(r, 0)$ tensor is an $(r, 0)$ tensor whose value is unaffected by the interchange (or *transposition*) of any two of its arguments, that is,

$$T(v_1, \dots, v_i, \dots, v_j, \dots, v_r) = T(v_1, \dots, v_j, \dots, v_i, \dots, v_r)$$

for any i and j . Symmetric $(0, r)$ tensors are defined similarly. You can easily check that the symmetric $(r, 0)$ and $(0, r)$ tensors each form vector spaces, denoted $S^r(V^*)$ and $S^r(V)$, respectively. For $T \in S^r(V^*)$, the symmetry condition implies that the components $T_{i_1 \dots i_r}$ are invariant under the transposition of any two indices, hence invariant under *any* rearrangement of the indices (since any rearrangement can be obtained via successive transpositions). Similar remarks apply, of course, to $S^r(V)$. Notice that for rank 2 tensors, the symmetry condition implies $T_{ij} = T_{ji}$ so that $[T]_{\mathcal{B}}$ for any \mathcal{B} is a symmetric matrix. Also note that it does not mean anything to say that a linear operator is symmetric, since a linear operator is a $(1, 1)$ tensor and there is no way of transposing the arguments. One might find that the matrix of a linear operator is symmetric in a certain basis, but this will not necessarily be true in other bases. If we have a metric to raise and lower indices then we can, of course, speak of symmetry by turning our linear operator into a $(2, 0)$ or $(0, 2)$ tensor.

Example 3.23 $S^2(\mathbb{R}^{2*})$

Consider the set $\{e^1 \otimes e^1, e^2 \otimes e^2, e^1 \otimes e^2 + e^2 \otimes e^1\} \subset S^2(\mathbb{R}^{2*})$ where $\{e^i\}_{i=1,2}$ is the standard dual basis. You can check that this set is linearly independent, and that any symmetric tensor can be written as

$$T = T_{11}e^1 \otimes e^1 + T_{22}e^2 \otimes e^2 + T_{12}(e^1 \otimes e^2 + e^2 \otimes e^1) \quad (3.74)$$

so this set is a basis for $S^2(\mathbb{R}^{2*})$, which is thus three-dimensional. In particular, the Euclidean metric g on \mathbb{R}^2 can be written as

$$g = e^1 \otimes e^1 + e^2 \otimes e^2$$

since $g_{11} = g_{22} = 1$ and $g_{12} = g_{21} = 0$. Note that g would *not* take this simple form in a non-orthonormal basis. \square

Exercise 3.25 Let $V = \mathbb{R}^n$ with the standard basis \mathcal{B} . Convince yourself that

$$[e_i \otimes e_j + e_j \otimes e_i]_{\mathcal{B}} = S_{ij}$$

where S_{ij} is the symmetric matrix defined in Example 2.8.

There are many symmetric tensors in physics, almost all of them of rank 2. Many of them we have met already: the Euclidean metric on \mathbb{R}^3 , the Minkowski metric on \mathbb{R}^4 , the moment of inertia tensor, and the Maxwell stress tensor. You should refer to the examples and check that these are all symmetric tensors. We have also met one class of higher rank symmetric tensors: the multipole moments.

Example 3.24 *Multipole moments and harmonic polynomials*

Recall from Example 3.4 that the scalar potential $\Phi(\mathbf{r})$ of a charge distribution $\rho(\mathbf{r}')$ localized around the origin in \mathbb{R}^3 can be expanded in a Taylor series in $1/r$ as

$$\Phi(\mathbf{r}) = \frac{1}{4\pi} \left[\frac{Q_0}{r} + \frac{Q_1(\mathbf{r})}{r^3} + \frac{1}{2!} \frac{Q_2(\mathbf{r}, \mathbf{r})}{r^5} + \frac{1}{3!} \frac{Q_3(\mathbf{r}, \mathbf{r}, \mathbf{r})}{r^7} + \dots \right]$$

where the Q_l are the symmetric rank l multipole moment tensors. Each symmetric tensor Q_l can be interpreted as a degree l polynomial f_l , just by evaluating on l copies of $\mathbf{r} = (x^1, x^2, x^3)$ as indicated:

$$f_l(\mathbf{r}) \equiv Q_l(\mathbf{r}, \dots, \mathbf{r}) = Q_{i_1 \dots i_l} x^{i_1} \dots x^{i_l}, \quad (3.75)$$

where the indices i_j are the usual component indices, *not* exponents. Note that the expression $x^{i_1} \dots x^{i_l}$ in the right hand side is invariant under any rearrangement of the indices i_j . This is because we fed in l copies of the *same* vector \mathbf{r} into Q_l . This fits in nicely with the symmetry of $Q_{i_1 \dots i_l}$. In fact, the above equation gives a one-to-one correspondence between l th rank symmetric tensors and degree l polynomials; we will not prove this correspondence here, but it should not be too hard to see that (3.75) turns any symmetric tensor into a polynomial, and that, conversely, any fixed degree polynomial can be written in the form of the right hand side of (3.75) with $Q_{i_1 \dots i_l}$ symmetric. This (roughly) explains why the multipole moments are symmetric tensors: the multipole moments are really just fixed degree polynomials, which in turn correspond to symmetric tensors.

What about the tracelessness of the Q_l , i.e. the fact that $\sum_k Q_{i_1 \dots k \dots k \dots i_n} = 0$? Well, $\Phi(\mathbf{r})$ obeys the Laplace equation $\Delta \Phi(\mathbf{r}) = 0$, which means that every term in the expansion must also obey the Laplace equation. Each term is of the form

$$\frac{f_l(\mathbf{r})}{r^{2l+1}},$$

and if we write the polynomial $f_l(\mathbf{r})$ as $r^l Y(\theta, \phi)$ then a quick computation shows that $Y(\theta, \phi)$ must be a spherical harmonic of degree l and hence f_l must be a harmonic polynomial! Expanding $f_l(\mathbf{r})$ in the form (3.75) and applying the Laplacian then shows that if f_l is harmonic, then Q_l must be traceless. \square

Exercise 3.26 What is the polynomial associated to the Euclidean metric tensor $g = \sum_{i=1}^3 e^i \otimes e^i$? What is the symmetric tensor in $S^3(\mathbb{R}^3)$ associated to the polynomial $x^2 y$?

Exercise 3.27 Substitute $f_l(\mathbf{r}) = r^l Y(\theta, \phi)$ into the equation

$$\Delta \left(\frac{f_l(\mathbf{r})}{r^{2l+1}} \right) = 0 \quad (3.76)$$

and show that $Y(\theta, \phi)$ must be a spherical harmonic of degree l . Then use (3.75) to show that if f_l is a harmonic polynomial, then the associated symmetric tensor Q_l must be traceless. If you have trouble showing that Q_l is traceless for arbitrary l , try starting with the $l = 2$ (quadrupole) and $l = 3$ (octopole) cases.

3.9 Antisymmetric Tensors

Now we turn to antisymmetric tensors. An *antisymmetric* (or *alternating*) $(r, 0)$ tensor is one whose value *changes sign* under transposition of any two of its arguments, i.e.

$$T(v_1, \dots, v_i, \dots, v_j, \dots, v_r) = -T(v_1, \dots, v_j, \dots, v_i, \dots, v_r). \quad (3.77)$$

Again, antisymmetric $(0, r)$ tensors are defined similarly and both sets form vector spaces, denoted $\Lambda^r V^*$ and $\Lambda^r V$ (for $r = 1$ we define $\Lambda^1 V^* = V^*$ and $\Lambda^1 V = V$). The following properties of antisymmetric tensors follow directly from (3.77); the first one is immediate, and the second two you will prove in Exercise 3.28 below:

- 1. $T(v_1, \dots, v_r) = 0$ if $v_i = v_j$ for any $i \neq j$
- \implies 2. $T(v_1, \dots, v_r) = 0$ if $\{v_1, \dots, v_r\}$ is linearly dependent
- \implies 3. If $\dim V = n$, then the only tensor in $\Lambda^r V^*$ and $\Lambda^r V$ for $r > n$ is the 0 tensor.

An important operation on antisymmetric tensors is the **wedge product**: Given $f, g \in V^*$ we define the wedge product of f and g , denoted $f \wedge g$, to be the antisymmetric $(2, 0)$ tensor defined by

$$f \wedge g \equiv f \otimes g - g \otimes f. \quad (3.78)$$

Note that $f \wedge g = -g \wedge f$, and that $f \wedge f = 0$. Expanding (3.78) in terms of the e^i gives

$$f \wedge g = f_i g_j (e^i \otimes e^j - e^j \otimes e^i) = f_i g_j e^i \wedge e^j \quad (3.79)$$

so that $\{e^i \wedge e^j\}_{i < j}$ spans all wedge products of dual vectors (note the “ $i < j$ ” stipulation, since $e^i \wedge e^j$ and $e^j \wedge e^i$ are not linearly independent). In fact, you can check that $\{e^i \wedge e^j\}_{i < j}$ is linearly independent and spans $\Lambda^2 V^*$, hence is a basis for $\Lambda^2 V^*$. The wedge product can be extended to r -fold products of dual vectors as follows: given r dual vectors f_1, \dots, f_r , we define their wedge product $f_1 \wedge \dots \wedge f_r$ to be the sum of all tensor products of the form $f_{i_1} \otimes \dots \otimes f_{i_r}$ where each term gets a $+$ or a $-$ sign depending on whether an odd or an even number¹⁵ of transpositions of the factors are necessary to obtain it from $f_1 \otimes \dots \otimes f_r$; if the number is odd the term is assigned -1 , if even a $+1$. Thus,

$$f_1 \wedge f_2 = f_1 \otimes f_2 - f_2 \otimes f_1 \quad (3.80)$$

$$\begin{aligned} f_1 \wedge f_2 \wedge f_3 &= f_1 \otimes f_2 \otimes f_3 + f_2 \otimes f_3 \otimes f_1 + f_3 \otimes f_1 \otimes f_2 \\ &\quad - f_3 \otimes f_2 \otimes f_1 - f_2 \otimes f_1 \otimes f_3 - f_1 \otimes f_3 \otimes f_2 \end{aligned} \quad (3.81)$$

and so on. You should convince yourself that $\{e^{i_1} \wedge \dots \wedge e^{i_r}\}_{i_1 < \dots < i_r}$ is a basis for $\Lambda^r V^*$. Note that this entire construction can be carried out for vectors as well as

¹⁵The number of transpositions required to get a given rearrangement is not unique, of course, but hopefully you can convince your self that it is always odd or always even. A rearrangement which always decomposes into an odd number of transpositions is an **odd rearrangement**, and **even rearrangements** are defined similarly. We will discuss this further in Chap. 4, specifically in Example 4.22.

dual vectors. Also note that all the comments about symmetry above Example 3.23 apply here as well.

Exercise 3.28 Let $T \in \Lambda^r V^*$. Show that if $\{v_1, \dots, v_r\}$ is a linearly dependent set then $T(v_1, \dots, v_r) = 0$. Use the same logic to show that if $\{f_1, \dots, f_r\} \subset V^*$ is linearly dependent, then $f_1 \wedge \dots \wedge f_r = 0$. If $\dim V = n$, show that any set of more than n vectors must be linearly dependent, so that $\Lambda^r V = \Lambda^r V^* = 0$ for $r > n$.

Exercise 3.29 Expand the $(2, 0)$ electromagnetic field tensor of (3.62) in the basis $\{e^i \wedge e^j\}$ where $i < j$ and $i, j = 1, 2, 3, 4$.

Exercise 3.30 Let $\dim V = n$. Show that the dimension of $\Lambda^r V^*$ and $\Lambda^r V$ is $\binom{n}{r} = \frac{n!}{(n-r)!r!}$.

Example 3.25 Identical particles

In quantum mechanics we often consider systems which contain *identical particles*, i.e. particles of the same mass, charge and spin. For instance, we might consider n non-interacting hydrogen atoms moving in a potential well, or the two electrons of the helium atom orbiting around the nucleus. In such cases we would assume that the total Hilbert space \mathcal{H}_{tot} would be just the n -fold tensor product of the single particle Hilbert space \mathcal{H} . It turns out, however, that nature does not work that way; for certain particles (known as **bosons**) only states in $S^n(\mathcal{H})$ are observed, while for other particles (known as **fermions**) only states in $\Lambda^n \mathcal{H}$ are observed. All known particles are either fermions or bosons. This restriction of the total Hilbert space to either $S^n(\mathcal{H})$ or $\Lambda^n \mathcal{H}$ is known as the **symmetrization postulate** and has far-reaching consequences. For instance, if we have two fermions, we cannot measure the same values for a complete set of quantum numbers for both particles, since then the state would have to include a term of the form $|\psi\rangle|\psi\rangle$ and thus could not belong to $\Lambda^2 \mathcal{H}$. This fact that two fermions cannot be in the same state is known as the **Pauli Exclusion Principle**. As another example, consider two identical spin 1/2 fermions fixed in space, so that $\mathcal{H}_{tot} = \Lambda^2 \mathbb{C}^2$. $\Lambda^2 \mathbb{C}^2$ is one-dimensional with basis vector

$$|0, 0\rangle = \left| \frac{1}{2} \right\rangle \left| -\frac{1}{2} \right\rangle - \left| -\frac{1}{2} \right\rangle \left| \frac{1}{2} \right\rangle$$

where we have used the notation of Example 3.21. If we measure \mathbf{S}^2 or S_z for this system we will get 0. This is in marked contrast to the case of two *distinguishable* spin 1/2 fermions; in this case the Hilbert space is $\mathbb{C}^2 \otimes \mathbb{C}^2$ and we have additional possible state kets

$$\begin{aligned} |1, 1\rangle &= \left| \frac{1}{2} \right\rangle \left| \frac{1}{2} \right\rangle \\ |1, 0\rangle &= \left| \frac{1}{2} \right\rangle \left| -\frac{1}{2} \right\rangle + \left| -\frac{1}{2} \right\rangle \left| \frac{1}{2} \right\rangle \\ |1, -1\rangle &= \left| -\frac{1}{2} \right\rangle \left| -\frac{1}{2} \right\rangle \end{aligned}$$

which yield nonzero values for \mathbf{S}^2 and S_z . □

The next two examples are a little more mathematical than physical but they are necessary for the discussion of pseudovectors below. Hopefully you will also find them of interest in their own right.

Example 3.26 *The Levi-Civita tensor*

Consider \mathbb{R}^n with the standard inner product. Let $\{e_i\}_{i=1,\dots,n}$ be an orthonormal basis for \mathbb{R}^n and consider the tensor

$$\boxed{\epsilon \equiv e^1 \wedge \dots \wedge e^n} \in \Lambda^n \mathbb{R}^{n*}.$$

You can easily check that

$$\epsilon_{i_1 \dots i_n} = \begin{cases} 0 & \text{if } \{i_1, \dots, i_n\} \text{ contains a repeated index} \\ -1 & \text{if } \{i_1, \dots, i_n\} \text{ is an odd rearrangement of } \{1, \dots, n\} \\ +1 & \text{if } \{i_1, \dots, i_n\} \text{ is an even rearrangement of } \{1, \dots, n\}. \end{cases}$$

For $n = 3$, you can also check (Exercise 3.31) that ϵ_{ijk} has the same values as the Levi-Civita symbol, and so ϵ here is an n -dimensional generalization of the three-dimensional Levi-Civita tensor we introduced in Example 3.2. As in that example, ϵ should be thought of as eating n vectors and spitting out the n -dimensional volume spanned by those vectors. This can be seen explicitly for $n = 2$ also; considering two vectors u and v in the (x, y) plane, we have

$$\begin{aligned} \epsilon(u, v) &= \epsilon_{ij} u^i v^j \\ &= u^x v^y - u^y v^x \\ &= (u \times v)^z \end{aligned}$$

and we know that this last expression can be interpreted as the (signed) area of the parallelogram spanned by u and v .

Finally, note that $\Lambda^n \mathbb{R}^{n*}$ is one-dimensional, and that ϵ is the basis for it described under (3.81).

You may object that our construction of ϵ seems to depend on a choice of metric and orthonormal basis. The former is true: ϵ *does* depend on the metric, and we make no apologies for that. As to whether it depends on a particular choice of orthonormal basis, we must do a little bit of investigating; this will require a brief detour into the subject of determinants.

Exercise 3.31 Check that the ϵ tensor on \mathbb{R}^3 satisfies

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } \{ijk\} \text{ is a cyclic permutation}^{16} \text{ of } \{1, 2, 3\} \\ -1 & \text{if } \{ijk\} \text{ is an anti-cyclic permutation of } \{1, 2, 3\} \\ 0 & \text{otherwise.} \end{cases}$$

Is it true for ϵ on \mathbb{R}^4 that $\epsilon_{ijkl} = 1$ if $\{ijkl\}$ is a cyclic permutation of $\{1, 2, 3, 4\}$?

¹⁶A **cyclic permutation** of $\{1, \dots, n\}$ is any rearrangement of $\{1, \dots, n\}$ obtained by successively moving numbers from the beginning of the sequence to the end. That is, $\{2, \dots, n, 1\}$, $\{3, \dots, n, 1, 2\}$, and so on are the cyclic permutations of $\{1, \dots, n\}$. **Anti-cyclic** permutations are cyclic permutations of $\{n, n-1, \dots, 1\}$.

Example 3.27 *The determinant*

You have doubtless encountered determinants before, and have probably seen them defined iteratively; that is, the determinant of a 2×2 square matrix A , denoted $|A|$ (or $\det A$), is defined to be

$$|A| \equiv A_{11}A_{22} - A_{12}A_{21} \quad (3.82)$$

and then the determinant of a 3×3 matrix B is defined in terms of this, i.e.

$$|B| \equiv B_{11} \begin{vmatrix} B_{22} & B_{23} \\ B_{32} & B_{33} \end{vmatrix} - B_{12} \begin{vmatrix} B_{21} & B_{23} \\ B_{31} & B_{33} \end{vmatrix} + B_{13} \begin{vmatrix} B_{21} & B_{22} \\ B_{31} & B_{32} \end{vmatrix}. \quad (3.83)$$

This expression is known as the **cofactor expansion** of the determinant, and is *not* unique; one can expand about any row (or column), not necessarily (B_{11}, B_{12}, B_{13}) .

In our treatment of the determinant we will take a somewhat more sophisticated approach.¹⁷ Take an $n \times n$ matrix A and consider its n columns as n column vectors in \mathbb{R}^n , so that the first column vector A_1 has i th component A_{i1} and so on. Then, constructing the ϵ tensor using the standard basis and inner product on \mathbb{R}^n , we define the **determinant of** A , denoted $|A|$ or $\det A$, to be

$$|A| \equiv \epsilon(A_1, \dots, A_n) \quad (3.84)$$

or in components

$$|A| = \sum_{i_1, \dots, i_n} \epsilon_{i_1, \dots, i_n} A_{i_1 1} \dots A_{i_n n}. \quad (3.85)$$

You should check explicitly that this definition reproduces (3.82) and (3.83) for $n = 2, 3$. You can also check in the Problems that many of the familiar properties of determinants (sign change under interchange of columns, invariance under addition of rows, factoring of scalars) follow quite naturally from the definition and the multilinearity and antisymmetry of ϵ .

Since the determinant is defined in terms of the epsilon tensor, which has an interpretation in terms of volume, this suggests that the determinant has an interpretation in terms of volume as well. Consider our matrix A as a linear operator on \mathbb{R}^n ; then A sends the standard orthonormal basis $\{e_1, \dots, e_n\}$ to a new, potentially non-orthonormal basis $\{Ae_1, \dots, Ae_n\}$. See Fig. 3.6. Then just as $\{e_1, \dots, e_n\}$ spans a regular n -cube whose volume is 1, the vectors $\{Ae_1, \dots, Ae_n\}$ span a skewed n -cube whose volume is given by

$$\begin{aligned} \epsilon(Ae_1, \dots, Ae_n) &= \epsilon\left(\sum_{i_1} A_{i_1 1} e_{i_1}, \dots, \sum_{i_n} A_{i_n n} e_{i_n}\right) \\ &= \sum_{i_1, \dots, i_n} A_{i_1 1} \dots A_{i_n n} \epsilon_{i_1, \dots, i_n} \\ &= \det A! \end{aligned}$$

¹⁷For a complete treatment, however, you should consult Hoffman and Kunze [10], Chap. 5.

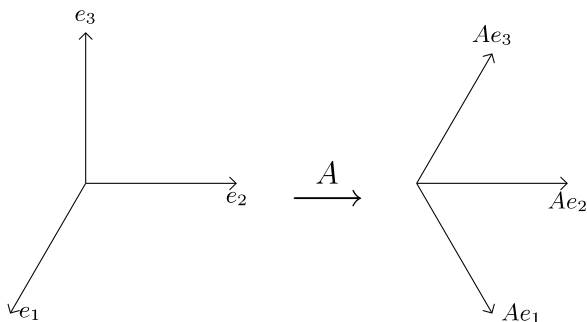


Fig. 3.6 The action of A on the standard basis in \mathbb{R}^3

Thus, **the determinant of a matrix is the (oriented) volume of the skew n -cube obtained by applying A to the standard n -cube!** You may have noticed that this volume can be negative, which is why we called the determinant an *oriented* (or *signed*) volume; the interpretation of this is given in the next example.

Example 3.28 *Orientations and the ϵ tensor*

Note This material is a bit abstract and may be skipped on a first reading.

With the determinant in hand we may now explore to what extent the definition of ϵ depends on our choice of orthonormal basis. Consider another orthonormal basis $\{e_{i'} = A_{i'}^j e_j\}$. If we define an ϵ' in terms of this basis, we find

$$\begin{aligned}
 \epsilon' &= e^{1'} \wedge \cdots \wedge e^{n'} \\
 &= A_{i_1}^{1'} \cdots A_{i_n}^{n'} e^{i_1} \wedge \cdots \wedge e^{i_n} \\
 &= A_{i_1}^{1'} \cdots A_{i_n}^{n'} \epsilon^{i_1 \dots i_n} e^1 \wedge \cdots \wedge e^n \\
 &= |A| \epsilon
 \end{aligned} \tag{3.86}$$

where in the third equality we used the fact that if $e^{i_1} \wedge \cdots \wedge e^{i_n}$ does not vanish it can always be rearranged to give $e^1 \wedge \cdots \wedge e^n$, and the resulting sign change if any is accounted for by the Levi-Civita symbol. Now since both $\{e_i\}$ and $\{e_{i'}\}$ are orthonormal bases, A must be an orthonormal matrix, so we can use the product rule for determinants $|AB| = |A||B|$ (see Problem 3.4 for a simple proof) and the fact that $|A^T| = |A|$ to get

$$1 = |I| = |AA^T| = |A||A^T| = |A|^2 \tag{3.87}$$

which implies $|A| = \pm 1$. Thus by (3.86) $\epsilon' = \epsilon$ if the two orthonormal bases used in their construction are related by an orthogonal transformation A with $|A| = 1$; such a transformation is called a **rotation**,¹⁸ and two bases related by a rotation, or by any

¹⁸No doubt you are used to thinking about a rotation as a transformation that preserves distances and fixes a line in space (the **axis of rotation**). This definition of a rotation is particular to \mathbb{R}^3 , since

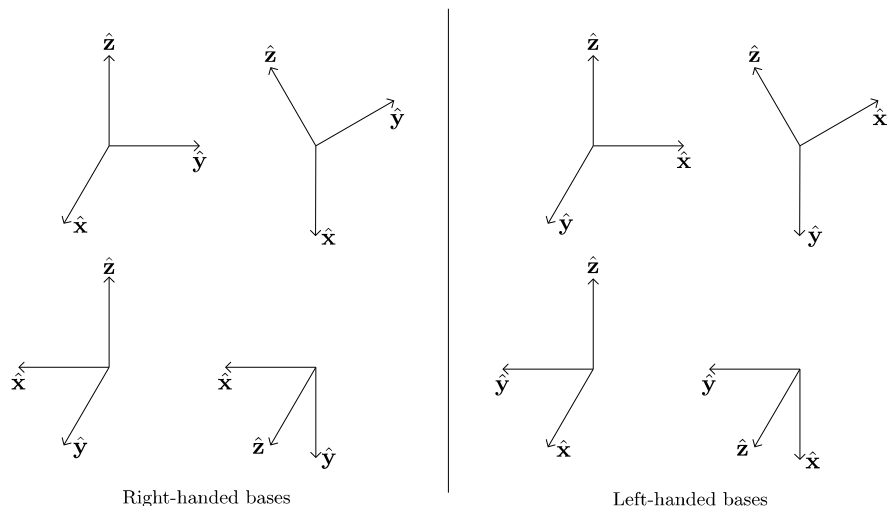


Fig. 3.7 The two orientations of \mathbb{R}^3 . The *upper-left most basis* is usually considered the standard basis

transformation with $|A| > 0$, are said to have the **same orientation**. If two bases are related by a basis transformation with $|A| < 0$ then the two bases are said to have the **opposite orientation**. We can then define an **orientation** as a maximal¹⁹ set of bases all having the same orientation, and you can show (see Problem 3.6) that \mathbb{R}^n has exactly two orientations. In \mathbb{R}^3 these two orientations are the right-handed bases and the left-handed bases, and are depicted schematically in Fig. 3.7. Thus we can say that ϵ does not depend on a particular choice of orthonormal basis, but it does depend on a metric and a choice of orientation, where the orientation chosen is the one determined by the standard basis.

The notion of orientation allows us to understand the interpretation of the determinant as an ‘oriented’ volume: the sign of the determinant just tells us whether or not the orientation of $\{Ae_i\}$ is the same as $\{e_i\}$. Also, for orientation-changing transformations on \mathbb{R}^3 one can show that A can be written as $A = A_0(-I)$, where A_0 is a rotation and $-I$ is referred to as the **inversion** transformation. The inversion transformation plays a key role in the next example.

Example 3.29 Pseudovectors in \mathbb{R}^3

Note All indices in this example refer to *orthonormal* bases. The calculations and results below do not apply to non-orthonormal bases.

even in \mathbb{R}^2 a rotation cannot be considered to be “about an axis” since $\hat{\mathbf{z}} \notin \mathbb{R}^2$. For the equivalence of our general definition and the more intuitive definition in \mathbb{R}^3 , see Goldstein [6].

¹⁹I.e. could not be made bigger.

A **pseudovector** (or **axial vector**) is a tensor on \mathbb{R}^3 whose components transform like vectors under rotations but do not change sign under inversion. Common examples of pseudovectors are the angular velocity vector ω , the magnetic field vector \mathbf{B} , as well as all cross products, such as the angular momentum vector $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. It turns out that pseudovectors like these are actually elements of $\Lambda^2\mathbb{R}^3$, which are known as **bivectors**.

To see the connection, consider the wedge product of two vectors $r, p \in \mathbb{R}^3$. This looks like

$$\begin{aligned} r \wedge p &= (r^1 e_1 + r^2 e_2 + r^3 e_3) \wedge (p^1 e_1 + p^2 e_2 + p^3 e_3) \\ &= (r^1 p^2 - r^2 p^1) e_1 \wedge e_2 + (r^3 p^1 - r^1 p^3) e_3 \wedge e_1 + (r^2 p^3 - r^3 p^2) e_2 \wedge e_3. \end{aligned}$$

This looks just like $r \times p$ if we make the identifications

$$\begin{aligned} e_1 \wedge e_2 &\longrightarrow e_3 \\ e_3 \wedge e_1 &\longrightarrow e_2 \\ e_2 \wedge e_3 &\longrightarrow e_1. \end{aligned} \tag{3.88}$$

In terms of matrices, this corresponds to the identification

$$\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{3.89}$$

This identification can be embodied in a one-to-one and onto map J from $\Lambda^2\mathbb{R}^3$ to \mathbb{R}^3 . If $\alpha \in \Lambda^2\mathbb{R}^3$ then we can expand it as

$$\alpha = \alpha^{23} e_2 \wedge e_3 + \alpha^{31} e_3 \wedge e_1 + \alpha^{12} e_1 \wedge e_2,$$

and we then define J in component as

$$\begin{aligned} J : \Lambda^2\mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ \alpha^{ij} &\mapsto (J(\alpha))^i \equiv \frac{1}{2} \epsilon^i_{jk} \alpha^{jk}. \end{aligned} \tag{3.90}$$

You will check below that this definition really does give the identifications written above. Note that J is essentially just a contraction with the epsilon tensor. With this, we see that $r \times p$ is really just $J(r \wedge p)$! Thus, **cross products are essentially just bivectors**.

Exercise 3.32 Check that J , as defined by (3.90), acts on basis vectors as in (3.88). Also check that when written in terms of matrices, J produces the map (3.89).

Now that we know how to identify bivectors and regular vectors, we must examine what it means for bivectors to transform ‘like’ vectors under rotations, but without a sign change under inversion. On the face of things, it seems like bivectors should transform very differently from vectors; after all, a bivector is a $(0, 2)$ tensor, and you can show that it has matrix transformation law

$$[\alpha]_{B'} = A[\alpha]_B A^T. \tag{3.91}$$

This looks very different from the transformation law for the associated vector $J(\alpha)$, which is just (cf. (3.29))

$$[J(\alpha)]_{\mathcal{B}'} = A[J(\alpha)]_{\mathcal{B}}. \quad (3.92)$$

In particular, the bivector α transforms with two copies of A , and the vector $J(\alpha)$ with just one. How could these transformation laws be ‘the same’? Well, remember that α is not any old $(0, 2)$ tensor, but an *antisymmetric* one, and for these a small miracle happens. This is best appreciated by considering an example. Let A be a rotation about the z -axis by an angle θ , so that

$$A = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then you can check that, with $\alpha^i \equiv (J(\alpha))^i$,

$$\begin{aligned} [\alpha]_{\mathcal{B}'} &= A[\alpha]_{\mathcal{B}}A^T \\ &= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \\ \alpha^3 & 0 & -\alpha^1 \\ -\alpha^2 & \alpha^1 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\alpha^3 & \alpha^2 \cos \theta + \alpha^1 \sin \theta \\ \alpha^3 & 0 & \alpha^2 \sin \theta - \alpha^1 \cos \theta \\ -\alpha^2 \cos \theta - \alpha^1 \sin \theta & -\alpha^2 \sin \theta + \alpha^1 \cos \theta & 0 \end{pmatrix} \end{aligned} \quad (3.93)$$

$$\xrightarrow{J} \begin{pmatrix} -\alpha^2 \sin \theta + \alpha^1 \cos \theta \\ \alpha^2 \cos \theta + \alpha^1 \sin \theta \\ \alpha^3 \end{pmatrix} \quad (3.94)$$

which is exactly a rotation about the z -axis of $J(\alpha)$! This seems to suggest that if we transform the components of $\alpha \in \Lambda^2 \mathbb{R}^3$ by a rotation *first* and then apply J , or apply J and *then* rotate the components, we get the same thing. In other words, **the map J commutes with rotations, and that is what it means for both bivectors and vectors to behave ‘the same’ under rotations.**

Exercise 3.33 Derive (3.91). You may need to consult Sect. 3.2.

Exercise 3.34 Verify (3.93) by performing the necessary matrix multiplication.

To *prove* that J commutes with *arbitrary* rotations (the example above just proved it for rotations about the z -axis), we need to show that

$$A^i_j \alpha^j = \frac{1}{2} \epsilon^{i' l'} A_m^{k'} A_n^{l'} \alpha^{mn} \quad (3.95)$$

where A is a rotation. On the left hand side J is applied first followed by a rotation, and on the right hand side the rotation is done first, followed by J .

We now compute

$$\begin{aligned}
 \frac{1}{2} \epsilon^{i' k' l'} A_m^{k'} A_n^{l'} \alpha^{mn} &= \frac{1}{2} \epsilon_{p' k' l'} \delta^{i' p'} A_m^{k'} A_n^{l'} \alpha^{mn} \\
 &= \frac{1}{2} \sum_q \epsilon_{p' k' l'} A_q^{i'} A_q^{p'} A_m^{k'} A_n^{l'} \alpha^{mn} \\
 &= \frac{1}{2} \sum_q \epsilon_{qmn} |A| A_q^{i'} \alpha^{mn} \\
 &= \frac{1}{2} |A| \epsilon_{mn}^q A_q^{i'} \alpha^{mn} \\
 &= |A| A_q^{i'} \alpha^q
 \end{aligned} \tag{3.96}$$

where in the second equality we used a variant of (3.32) which comes from writing out $AA^T = I$ in components, in the third equality we used the easily verified fact that $\epsilon_{p' k' l'} A_q^{p'} A_m^{k'} A_n^{l'} = |A| \epsilon_{qmn}$, and in the fourth equality we raised an index to resume use of Einstein summation convention and were able to do so because covariant and contravariant components are equal in orthonormal bases. Now, for *rotations* $|A| = 1$ so in this case (3.96) and (3.95) are identical and the components of $J(\alpha)$ transform like the components of a vector. For *inversion*, however, $|A| = |-I| = -1$ so (3.96) tells us that the components of $J(\alpha)$ do *not* change sign under inversion, as those of an ordinary vector would. Another way to see this is given in the exercise below.

Exercise 3.35 Use (3.91) to show that the components of $J(\alpha)$ do not change sign under inversion.

We have thus shown that pseudovectors are bivectors, since bivectors transform like vectors under rotation but do not change sign under inversion. We have also seen that cross products are very naturally interpreted as bivectors. There are other pseudovectors lying around, though, that do not naturally arise as cross products. For instance, what about the angular velocity vector ω ?

Example 3.30 *The angular velocity vector*

The angular velocity vector ω is usually introduced in the context of rigid body rotations. One usually fixes the center of mass of the body, and then the velocity \mathbf{v} of a point of the rigid body is given by

$$\mathbf{v} = \omega \times \mathbf{r} \tag{3.97}$$

where \mathbf{r} is the position vector of the point as measured from the center of mass. The derivation of this equation usually involves consideration of an angle and axis of rotation, and from these considerations one can argue that ω is a pseudovector. Here we will take a different approach in which ω will appear first as an antisymmetric matrix, making the bivector nature of ω manifest.

Let K and K' be two orthonormal bases for \mathbb{R}^3 as in Example 2.11, with K time-dependent. One should think of K as being attached to the rotating rigid body,

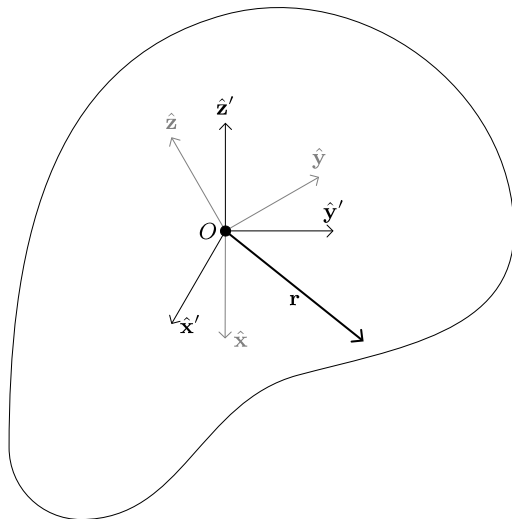


Fig. 3.8 Our rigid body with fixed space frame K' and body frame K in gray

whereas K' is fixed. We will refer to K as the **body** frame and K' as the **space** frame. Both frames have their origin at the center of mass of the rigid body. This is depicted in Fig. 3.8.

Now let \mathbf{r} represent a point of the rigid body; then $[\mathbf{r}]_K$ will be its coordinates in the body frame, and $[\mathbf{r}]_{K'}$ its coordinates in the space frame. Let A be the (time-dependent) matrix of the basis transformation taking K' to K , so that

$$[\mathbf{r}]_{K'} = A[\mathbf{r}]_K. \quad (3.98)$$

We would now like to calculate the velocity $[\mathbf{v}]_{K'}$ of our point in the rigid body, as seen in the space frame, and compare it to (3.97). This is given by just differentiating $[\mathbf{r}]_{K'}$:

$$\begin{aligned} [\mathbf{v}]_{K'} &= \frac{d}{dt}[\mathbf{r}]_{K'} \\ &= \frac{d}{dt}A[\mathbf{r}]_K \quad \text{by (3.98)} \\ &= \frac{dA}{dt}[\mathbf{r}]_K \quad \text{since } [\mathbf{r}]_K \text{ is constant} \\ &= \frac{dA}{dt}A^{-1}[\mathbf{r}]_{K'}. \end{aligned} \quad (3.99)$$

So far this does not really look like (3.97). What to do? First, observe that $\frac{dA}{dt}A^{-1}$ is actually an antisymmetric matrix:

$$\begin{aligned} 0 &= \frac{d}{dt}(I) \\ &= \frac{d}{dt}(AA^T) \end{aligned}$$

$$\begin{aligned}
&= \frac{dA}{dt} A^T + A \frac{dA^T}{dt} \\
&= \frac{dA}{dt} A^T + \left(\frac{dA}{dt} A^T \right)^T.
\end{aligned} \tag{3.100}$$

We can then define an **angular velocity bivector** $\tilde{\omega}$ whose components in the space frame are given by

$$[\tilde{\omega}]_{K'} = \frac{dA}{dt} A^{-1}. \tag{3.101}$$

Then we simply define the **angular velocity vector** ω to be

$$\omega \equiv J(\tilde{\omega}). \tag{3.102}$$

Note that ω is, in general, time-dependent. It follows from this definition that

$$\frac{dA}{dt} A^{-1} = \begin{pmatrix} 0 & -\omega^{3'} & \omega^{2'} \\ \omega^{3'} & 0 & -\omega^{1'} \\ -\omega^{2'} & \omega^{1'} & 0 \end{pmatrix}$$

(we use primed indices since we are working with the K' components), so then

$$\begin{aligned}
\frac{dA}{dt} A^{-1} [\mathbf{r}]_{K'} &= \begin{pmatrix} 0 & -\omega^{3'} & \omega^{2'} \\ \omega^{3'} & 0 & -\omega^{1'} \\ -\omega^{2'} & \omega^{1'} & 0 \end{pmatrix} \begin{pmatrix} x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} \\
&= \begin{pmatrix} \omega^{2'} x^{3'} - \omega^{3'} x^{2'} \\ \omega^{3'} x^{1'} - \omega^{1'} x^{3'} \\ \omega^{1'} x^{2'} - \omega^{2'} x^{1'} \end{pmatrix} \\
&= [\omega \times \mathbf{r}]_{K'}!
\end{aligned}$$

Combining this with (3.99), we then have

$$[\mathbf{v}]_{K'} = [\omega \times \mathbf{r}]_{K'} \tag{3.103}$$

which is just (3.97) written in the space frame! Thus the ‘pseudovector’ ω can be viewed as nothing more than the vector associated to the antisymmetric matrix $\frac{dA}{dt} A^{-1}$, and multiplication by this is just the cross product with ω ! \square

Exercise 3.36 Use (3.91) to show that the bivector $\tilde{\omega}$ in the body frame is

$$[\tilde{\omega}]_K = A^{-1} \frac{dA}{dt}.$$

Combine this with (3.99) to show that (3.97) is true in the body frame as well.

In the last example we saw that to any time-dependent rotation matrix A we could associate an antisymmetric matrix $\frac{dA}{dt} A^{-1}$, which we can identify with the angular velocity vector which represents ‘infinitesimal’ rotations. This association between finite transformations and their infinitesimal versions, which in the case of

rotations takes us from orthogonal matrices to antisymmetric matrices, is precisely the relationship between a *Lie group* and its *Lie algebra*. We turn our attention to these objects in the next part of this book.

3.10 Problems

3.1 In this problem we explore the properties of $n \times n$ orthogonal matrices. This is the set of real invertible matrices A satisfying $A^T = A^{-1}$, and is denoted $O(n)$.

- (a) Is $O(n)$ a subspace of $M_n(\mathbb{R})$?
- (b) Show that the product of two orthogonal matrices is again orthogonal, that the inverse of an orthogonal matrix is again orthogonal, and that the identity matrix is orthogonal. These properties show that $O(n)$ is a **group**, i.e. a set with an associative multiplication operation and identity element such that the set is closed under multiplication and every element has a multiplicative inverse. Groups are the subject of Chap. 4.
- (c) Show that the columns of an orthogonal matrix A , viewed as vectors in \mathbb{R}^n , are mutually orthogonal under the usual inner product. Show the same for the rows. Show that for an active transformation, i.e.

$$[e_{i'}]_{\mathcal{B}} = A[e_i]_{\mathcal{B}}$$

where $\mathcal{B} = \{e_i\}_{i=1,\dots,n}$ so that

$$[e_i]_{\mathcal{B}}^T = (0, \dots, \underbrace{1}_{i\text{th slot}}, \dots, 0),$$

the columns of A are the $[e_{i'}]_{\mathcal{B}}$. In other words, the components of the *new* basis vectors in the *old* basis are just the columns of A . This also shows that for the corresponding passive transformation, in which

$$[e_i]_{\mathcal{B}'} = A^{-1}[e_i]_{\mathcal{B}}$$

the columns of A^{-1} are the components of the *old* basis vectors in the *new* basis.

- (d) Show that the orthogonal matrices A with $|A| = 1$, the rotations, form a subgroup unto themselves, denoted $SO(n)$. Do the matrices with $|A| = -1$ also form a subgroup?

3.2 In this problem we will compute the dimension of the space of $(0, r)$ symmetric tensors $S^r(V)$. This is slightly more difficult to compute than the dimension of the space of $(0, r)$ antisymmetric tensors $\Lambda^r V$, which was Exercise 3.30.

- (a) Let $\dim V = n$ and $\{e_i\}_{i=1,\dots,n}$ be a basis for V . Argue that $\dim S^r(V)$ is given by the number of ways you can choose r (possibly repeated) vectors from the basis $\{e_i\}_{i=1,\dots,n}$.

- (b) We have now reduced the problem to a combinatorics problem: how many ways can you choose r objects from a set of n objects, where any object can be chosen more than once? The answer is

$$\dim S^r(V) = \binom{n+r-1}{n-1} \equiv \frac{(n+r-1)!}{r!(n-1)!}. \quad (3.104)$$

Try to derive this on your own. If you need help, the solution to this is known as the “stars and bars” or “balls and walls” method; you can also refer to Sternberg [16], Chap. 5.

3.3 Prove the following basic properties of the determinant directly from the definition (3.84). We will restrict our discussion to operations with columns, though it can be shown that all the corresponding statements for rows are true as well.

- (a) Any matrix with a column of zeros has $|A| = 0$.
- (b) Multiplying a column by a scalar c multiplies the whole determinant by c .
- (c) The determinant changes sign under interchange of any two columns.
- (d) Adding two columns together, i.e. sending $A_i \rightarrow A_i + A_j$ for any i and j , does not change the value of the determinant.

3.4 One can extend the definition of determinants from matrices to more general linear operators as follows: We know that a linear operator T on a vector space V (equipped with an inner product and orthonormal basis $\{e_i\}_{i=1,\dots,n}$) can be extended to an operator on the p -fold tensor product $\mathcal{T}_p^0(V)$ by

$$T(v_1 \otimes \cdots \otimes v_p) = (Tv_1) \otimes \cdots \otimes (Tv_p).$$

Since $\Lambda^n V \subset \mathcal{T}_n^0(V)$, the action of T then extends to $\Lambda^n V$ similarly by

$$T(v_1 \wedge \cdots \wedge v_n) = (Tv_1) \wedge \cdots \wedge (Tv_n).$$

Consider then the action of T on the contravariant version of ϵ , the tensor $\tilde{\epsilon} \equiv e_1 \wedge \cdots \wedge e_n$. We know from Exercise 3.30 that $\Lambda^n V$ is one-dimensional, so that $T(\tilde{\epsilon}) = (Te_1) \wedge \cdots \wedge (Te_n)$ is proportional to $\tilde{\epsilon}$, and we define the **determinant** of T to be this proportionality constant, so that

$$(Te_1) \wedge \cdots \wedge (Te_n) \equiv |T| e_1 \wedge \cdots \wedge e_n. \quad (3.105)$$

- (a) Show by expanding the left hand side of (3.105) in components that this more general definition reduces to the old one of (3.85) in the case of $V = \mathbb{R}^n$.
- (b) Use this definition of the determinant to show that for two linear operators B and C on V ,

$$|BC| = |B||C|.$$

In particular, this result holds when B and C are square matrices.

- (c) Use (b) to show that the determinant of a matrix is invariant under similarity transformations (see Example 3.8). Conclude that we could have defined the determinant of a linear operator T as the determinant of its matrix in *any* basis.

3.5 Let V be a vector space with an inner product and orthonormal basis $\{e_i\}_{i=1,\dots,n}$. Prove that a linear operator T is invertible if and only if $|T| \neq 0$, as follows:

- (a) Show that T is invertible if and *only if* $\{T(e_i)\}_{i=1,\dots,n}$ is a linearly independent set (see Exercise 2.7 for the ‘if’ part of the statement).
- (b) Show that $|T| \neq 0$ if and only if $\{T(e_i)\}_{i=1,\dots,n}$ is a linearly independent set.
- (c) This is not a problem, just a comment. In Example 3.27 we interpreted the determinant of a matrix A as the oriented volume of the n -cube determined by $\{Ae_i\}$. As you just showed, if A is not invertible then the Ae_i are linearly dependent, hence span a space of dimension less than n and thus yield n -dimensional volume 0. Thus, the geometrical picture is consistent with the results you just obtained!

3.6 Let \mathcal{B} be the standard basis for \mathbb{R}^n , \mathcal{O} the set of all bases related to \mathcal{B} by a basis transformation with $|A| > 0$, and \mathcal{O}' the set of all bases related to \mathcal{B} by a transformation with $|A| < 0$.

- (a) Using what we have learned in the preceding problems, show that a basis transformation matrix A cannot have $|A| = 0$.
- (b) \mathcal{O} is by definition an orientation. Show that \mathcal{O}' is also an orientation, and conclude that \mathbb{R}^n has exactly two orientations. Note that both \mathcal{O} and \mathcal{O}' contain orthonormal and non-orthonormal bases.
- (c) For what n is $A = -I$ an orientation-changing transformation?

Part II

Group Theory

Chapter 4

Groups, Lie Groups, and Lie Algebras

In physics we are often interested in how a particular object behaves under a particular set of transformations. For instance, one often reads that under rotations a dipole moment transforms like a vector and a quadrupole moment like a second rank tensor, and that electric and magnetic fields transform independently like vectors under rotations but collectively like a second rank antisymmetric tensor under Lorentz transformations. Similarly, in quantum mechanics one is often interested in the “spin” of a ket (which specifies how it transforms under rotations), or its behavior under the time-reversal or space inversion (parity) transformations. This knowledge is particularly useful as it leads to the many famous “selection rules” which greatly simplify evaluation of matrix elements. Transformations are also crucial in quantum mechanics because, as we will see, all physical observables can be considered as “infinitesimal generators” of particular transformations; for example, the angular momentum operators “generate” rotations (as we discussed briefly in Problem 2.3) and the momentum operator “generates” translations.

Like tensors, this material is often treated in a somewhat ad hoc way which facilitates computation but obscures the underlying mathematical structures. These underlying structures are known to mathematicians as *group theory*, *Lie theory* and *representation theory*, and are known collectively to physicists as just “group theory”. Our aim in this second part of the book is to present the basic facts of this theory as well as its manifold applications to physics, both to clarify and unify the diverse phenomena in physics which it underlies and also to provide a nice application of what we have learned about tensors.

Before we discuss how particular objects transform, however, we must discuss the transformations themselves, in both their ‘finite’ and ‘infinitesimal’ form. That discussion is the subject of the present chapter. We will begin with a discussion of the ‘finite’ transformations, all of which share a few common properties: First, the performance of two successive transformations is always equivalent to the performance of a single, third transformation (just think of rotations and how any two successive rotations about two axes can be considered as a single rotation about a third axis). Second, every transformation has an inverse which undoes it (in the case of rotations, the inverse to any given rotation is a rotation about the same axis in the

opposite direction). Sets of transformations like these occur so often in mathematics and physics that they are given a name: *groups*.

4.1 Groups—Definition and Examples

The definition of a group that we are about to give may appear somewhat abstract, but is just meant to embody the most important properties of sets of transformations. After giving the definition and establishing some basic properties, we will proceed shortly to some concrete examples.

A **group** is a set G together with a ‘multiplication’ operation, denoted \cdot , that satisfies the following axioms:

1. (Closure) $g, h \in G$ implies $g \cdot h \in G$.
2. (Associativity) For $g, h, k \in G$, $g \cdot (h \cdot k) = (g \cdot h) \cdot k$.
3. (Existence of the identity) There exists an element $e \in G$ such that $g \cdot e = e \cdot g = g$ $\forall g \in G$.
4. (Existence of inverses) $\forall g \in G$ there exists an element $h \in G$ such that $g \cdot h = h \cdot g = e$.

If we think of a group as a set of transformations, as we usually do in physics, then the multiplication operation is obviously just composition; for instance, if R and S are three-dimensional rotations, then $R \cdot S$ is just S followed by R . Note that we do not necessarily have $R \cdot S = S \cdot R$ for all rotations R and S ; in cases such as this, G is said to be **non-commutative** (or **non-abelian**). If we did have $S \cdot R = R \cdot S$ for all $R, S \in G$, then we would say that G is **commutative** (or **abelian**).

There are several important properties of groups that follow almost immediately from the definition. Firstly, the *identity is unique*, for if e and f are both elements satisfying axiom 3 then we have

$$\begin{aligned} e &= e \cdot f \quad \text{since } f \text{ is an identity} \\ &= f \quad \text{since } e \text{ is an identity.} \end{aligned}$$

Secondly, *inverses are unique*: Let $g \in G$ and let h and k both be inverses of g . Then

$$g \cdot h = e \tag{4.1}$$

so multiplying both sides on the left by k gives

$$\begin{aligned} k \cdot (g \cdot h) &= k \cdot e, \\ (k \cdot g) \cdot h &= k \quad \text{by associativity,} \\ e \cdot h &= k \quad \text{since } k \text{ is an inverse of } g, \\ h &= k. \end{aligned}$$

We henceforth denote the unique inverse of an element g as g^{-1} .

Thirdly, if $g \in G$ and h is merely a *right* inverse for g , i.e.

$$g \cdot h = e, \quad (4.2)$$

then h is also a *left* inverse for g and is hence the unique inverse g^{-1} . This is seen as follows:

$$\begin{aligned} h \cdot g &= (g^{-1} \cdot g) \cdot (h \cdot g) \\ &= (g^{-1} \cdot (g \cdot h)) \cdot g \quad \text{by associativity} \\ &= (g^{-1} \cdot e) \cdot g \quad \text{by (4.2)} \\ &= g^{-1} \cdot g \\ &= e, \end{aligned}$$

so $h = g^{-1}$.

The last few properties concern inverses and can be verified immediately:

$$\begin{aligned} (g^{-1})^{-1} &= g \\ (g \cdot h)^{-1} &= h^{-1} \cdot g^{-1} \\ e^{-1} &= e. \end{aligned}$$

Exercise 4.1 Prove the *cancelation laws* for groups, i.e. that

$$\begin{aligned} g_1 \cdot h = g_2 \cdot h &\implies g_1 = g_2 \\ h \cdot g_1 = h \cdot g_2 &\implies g_1 = g_2. \end{aligned}$$

Before we get to some examples, we should note that properties 2 and 3 in the definition above are usually obviously satisfied and one rarely needs to check them explicitly. The important thing in showing that a set is a group is verifying that it is closed under multiplication and contains all its inverses. Also, as a matter of notation, from now on we will usually omit the \cdot when writing a product, and simply write gh for $g \cdot h$.

Example 4.1 \mathbb{R} : *The real numbers as an additive group*

Consider the real numbers \mathbb{R} with the group ‘multiplication’ operation given by regular addition, i.e.

$$x \cdot y \equiv x + y, \quad x, y \in \mathbb{R}. \quad (4.3)$$

It may seem counterintuitive to define ‘multiplication’ as addition, but the definition of a group is rather abstract so there is nothing that prevents us from doing this, and this point of view will turn out to be useful. With addition as the product, \mathbb{R} becomes an abelian group: The first axiom to verify is closure, and this is satisfied since the sum of two real numbers is always a real number. The associativity axiom is also satisfied, since it is a fundamental property of real numbers that addition is associative. The third axiom, dictating the existence of the identity, is satisfied since $0 \in \mathbb{R}$ fits the bill. The fourth axiom, which dictates the existence of inverses, is

satisfied since for any $x \in \mathbb{R}$, $-x$ is its (additive) inverse. Thus \mathbb{R} is a group under addition, and is in fact an abelian group since $x + y = y + x$, $\forall x, y \in \mathbb{R}$.

Note that \mathbb{R} is *not* a group under regular multiplication, since 0 has no multiplicative inverse. If we remove 0, though, then we do get a group, the multiplicative group of nonzero real numbers, denoted \mathbb{R}^* . We leave it to you to verify that \mathbb{R}^* is a group. You should also verify that this entire discussion goes through for \mathbb{C} as well, so that \mathbb{C} under addition and $\mathbb{C}^* \equiv \mathbb{C} \setminus \{0\}$ under multiplication are both abelian groups.

Example 4.2 Vector spaces as additive groups

The previous example of \mathbb{R} and \mathbb{C} as additive groups can be generalized to the case of vector spaces, which are abelian groups under vector space addition. The group axioms follow directly from the vector space axioms, as you should check, with 0 as the identity. While viewing vector spaces as additive groups means we ignore the crucial feature of scalar multiplication, we will see that this perspective will occasionally prove useful.

Example 4.3 $GL(V)$, $GL(n, \mathbb{R})$, and $GL(n, \mathbb{C})$: The general linear groups

The **general linear group** of a vector space V , denoted $GL(V)$, is defined to be the subset of $\mathcal{L}(V)$ consisting of all *invertible* linear operators on V . We can easily verify that $GL(V)$ is a group: to verify closure, note that for any $T, U \in GL(V)$, TU is linear and $(TU)^{-1} = U^{-1}T^{-1}$, so TU is invertible. To verify associativity, note that for any $T, U, V \in GL(V)$ and $v \in V$, we have

$$(T(UV))(v) = T(U(V(v))) = ((TU)V)(v) \quad (4.4)$$

(careful unraveling the meaning of the parentheses!) so that $T(UV) = (TU)V$. To verify the existence of the identity, just note that I is invertible and linear, hence in $GL(V)$. To verify the existence of inverses, note that for any $T \in GL(V)$, T^{-1} exists and is invertible and linear, hence is in $GL(V)$ also. Thus $GL(V)$ is a group.

Let V have scalar field C and dimension n . If we pick a basis for V , then for each $T \in GL(V)$ we get an invertible matrix $[T] \in M_n(C)$. Just as all the invertible $T \in \mathcal{L}(V)$ form a group, so do the corresponding invertible matrices in $M_n(C)$; this group is denoted as $GL(n, C)$, and the group axioms can be readily verified for it.¹ When $C = \mathbb{R}$ we get $GL(n, \mathbb{R})$, the **real general linear group in n dimensions**, and when $C = \mathbb{C}$ we get $GL(n, \mathbb{C})$, the **complex general linear group in n dimensions**. \square

While neither $GL(V)$, $GL(n, \mathbb{R})$, nor $GL(n, \mathbb{C})$ occur explicitly very often in physics, they have many important **subgroups**, i.e. subsets which themselves are groups. The most important of these arise when we have a vector space V equipped

¹You may recall having met $GL(n, \mathbb{R})$ at the end of Sect. 2.1. There we asked why it is not a vector space, and now we know—it is more properly thought of as a group!

with a non-degenerate Hermitian form $(\cdot | \cdot)$. In this case, we can consider the set of **isometries** $\text{Isom}(V)$, consisting of those operators T which ‘preserve’ $(\cdot | \cdot)$ in the sense that

$$(Tv|Tw) = (v|w) \quad \forall v, w \in V. \quad (4.5)$$

If $(\cdot | \cdot)$ can be interpreted as giving the ‘length’ of vectors, then an isometry T can be thought of as an operator that preserves lengths.² Note that any such T is invertible (why?), hence $\text{Isom}(V) \subset GL(V)$. $\text{Isom}(V)$ is in fact a subgroup of $GL(V)$, as we will now verify. First off, for any $T, U \in \text{Isom}(V)$,

$$\begin{aligned} ((TU)v|(TU)w) &= (T(Uv)|T(Uw)) \\ &= (Uv|Uw) \quad \text{since } T \text{ is an isometry} \\ &= (v|w) \quad \text{since } U \text{ is an isometry} \end{aligned}$$

so TU is an isometry as well, hence $\text{Isom}(V)$ is closed under multiplication. As for associativity, this axiom is automatically satisfied since $\text{Isom}(V) \subset GL(V)$ and multiplication in $GL(V)$ is associative, as we proved above. As for the existence of the identity, just note that the identity operator I is trivially an isometry. To verify the existence of inverses, note that T^{-1} exists and is an isometry since

$$\begin{aligned} (T^{-1}v|T^{-1}w) &= (TT^{-1}v|TT^{-1}w) \quad \text{since } T \in \text{Isom}(V) \\ &= (v|w). \end{aligned}$$

Thus $\text{Isom}(V)$ is a group. Why is it of interest? Well, as we will show in the next few examples, the matrix representations of $\text{Isom}(V)$ actually turn out to be the orthogonal matrices, the unitary matrices, and the Lorentz transformations, depending on whether or not V is real or complex and whether or not $(\cdot | \cdot)$ is positive-definite. Our discussion here shows³ that all of these sets of matrices are groups, and that they can all be thought of as representing linear operators which preserve the relevant non-degenerate Hermitian form.

Example 4.4 The orthogonal group $O(n)$

Let V be an n -dimensional real inner product space. The isometries of V can be thought of as operators which preserve lengths *and* angles, since the formula

$$\cos \theta = \frac{(v|w)}{\|v\|\|w\|}$$

for the angle between v and w is defined purely in terms of the inner product. Now, if T is an isometry and we write out $(v, w) = (Tv|Tw)$ in components referred to an orthonormal basis \mathcal{B} , we find that

²Hence the term ‘iso-metry’ = ‘same length’.

³We are glossing over some subtleties with this claim. See Example 4.17 for the full story.

$$\begin{aligned}
[v]^T[w] &= (v|w) \\
&= (Tv|Tw) \\
&= \delta_{ij}(Tv)^i(Tv)^j \\
&= \delta_{ij}T^i_k v^k T^j_l w^l \\
&= \sum_i v^k T^i_k T^i_l w^l \\
&= [v]^T [T]^T [T][w] \quad \forall v, w, \in V.
\end{aligned} \tag{4.6}$$

As you will show in Exercise 4.2, this is true if and only if $[T]^T [T] = I$, or

$$[T]^T = [T]^{-1}. \tag{4.7}$$

This is the familiar orthogonality condition, and the set of all orthogonal matrices (which, as we know from the above discussion and from Problem 3.1, form a group) is known as the **orthogonal group** $O(n)$. We will consider $O(n)$ in detail for $n = 2$ and 3 in the next section.

Exercise 4.2 Show that

$$[v]^T [w] = [v]^T [T]^T [T][w] \quad \forall v, w \in V \tag{4.8}$$

if and only if $[T]^T [T] = I$. One direction is easy; for the other, let $v = e_i$, $w = e_j$ where $\{e_i\}_{i=1, \dots, n}$ is orthonormal.

Exercise 4.3 Verify directly that $O(n)$ is a group, by using the defining condition (4.7). This is the same as Problem 3.1(b).

Example 4.5 The unitary group $U(n)$

Now let V be a complex inner product space. Recall from Problem 2.6 that we can define the **adjoint** T^\dagger of a linear operator T by the equation

$$\boxed{(T^\dagger v|w) \equiv (v|Tw)}. \tag{4.9}$$

If T is an isometry, then we can characterize it in terms of its adjoint, as follows: first, we have

$$\begin{aligned}
(v|w) &= (Tv|Tw) \\
&= (T^\dagger Tv|w) \quad \forall v, w \in V.
\end{aligned}$$

Calculations identical to those of Exercise 4.2 then show that this can be true if and only if $T^\dagger T = T T^\dagger = I$, which is equivalent to the more familiar condition

$$\boxed{T^\dagger = T^{-1}}. \tag{4.10}$$

Such an operator is said to be **unitary**. Thus every isometry of a complex inner product space is unitary, and vice versa. Now, if V has dimension n and we choose an orthonormal basis for V , then we have $[T^\dagger] = [T]^\dagger$ (cf. Problem 2.4), and so

(4.10) implies

$$[T]^\dagger = [T]^{-1}. \quad (4.11)$$

Thus, in an orthonormal basis, a unitary operator is represented by a unitary matrix! (Note that this is NOT necessarily true in a non-orthonormal basis.) By the discussion preceding Example 4.4, the set of all unitary matrices forms a group, denoted $U(n)$. We will not discuss $U(n)$ in depth in this text, but we will discuss one of its cousins, $SU(2)$ (to be defined below), extensively.

Note that there is nothing in the above discussion that *requires* V to be complex, so we can actually use the same definitions (of adjoints and unitarity) to define unitary operators on *any* inner product space, real or complex. Thus, a unitary operator is just an isometry of a real or complex inner product space.⁴ In the case of a real vector space, the unitary matrix condition (4.11) reduces to the orthogonality condition (4.7), as you might expect.

Exercise 4.4 Verify directly that $U(n)$ is a group, using the defining condition (4.11).

Example 4.6 *The Lorentz group* $O(n-1, 1)$

Now let V be a real vector space with a **Minkowski metric** η , which is defined, as in Example 2.19, as a symmetric, non-degenerate $(2, 0)$ tensor whose matrix in an orthonormal basis has the form

$$[\eta] = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \cdots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix} \quad (4.12)$$

with zeros on all the off-diagonals. This is to be compared with (2.25), which is just (4.12) with $n = 4$. Now, since η is a non-degenerate Hermitian form, we can consider its group of isometries. If $T \in \text{Isom}(V)$, then in analogy to the computation leading to (4.6), we have (in an arbitrary basis \mathcal{B}),

$$\begin{aligned} [v]^T [\eta] [w] &= \eta(v, w) \\ &= \eta(Tv, Tw) \\ &= [v]^T [T]^T [\eta] [T] [w] \quad \forall v, w \in V. \end{aligned} \quad (4.13)$$

Again, the same argument as you used in Exercise 4.2 shows that the above holds if and only if

$$[T]^T [\eta] [T] = [\eta] \quad (4.14)$$

⁴In fact, the only reason for speaking of both “isometries” and “unitary operators” is that unitary operators act solely on inner product spaces, whereas isometries can act on spaces with non-degenerate Hermitian forms that are not necessarily positive-definite, such as \mathbb{R}^4 with the Minkowski metric.

which in components reads

$$T_\mu{}^\rho T_\nu{}^\sigma \eta_{\rho\sigma} = \eta_{\mu\nu}. \quad (4.15)$$

If \mathcal{B} is orthonormal, (4.14) becomes

$$[T]^T \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cdots & \\ & & & 1 \\ & & & & -1 \end{pmatrix} [T] = [\eta] \quad (4.16)$$

which you will recognize from (3.37) as the definition of a Lorentz transformation, though now we are working in an arbitrary dimension n rather than just dimension four. We thus see that the set of all Lorentz transformations forms a group, known as the **Lorentz group** and denoted by $O(n-1, 1)$ (the notation just refers to the number of positive and negative 1s present in the matrix form of η given in (4.12)). The Lorentz transformations lie at the heart of special relativity, and we will take a close look at these matrices for $n = 4$ in the next section. \square

Exercise 4.5 Verify directly that $O(n-1, 1)$ is group, using the defining condition (4.14).

Aside You may recall that we originally defined orthogonal matrices, unitary matrices, and Lorentz transformations as those matrices which implement a basis change from one orthonormal basis to another (on vector spaces with real inner products, Hermitian inner products, and Minkowski metrics, respectively). In the preceding examples, however, we have seen that these matrices can alternatively be defined as those which represent (in an orthonormal basis) operators which preserve a non-degenerate Hermitian form. These two definitions correspond to the active and passive viewpoints of transformations: our first definition of these matrices (as those which implement orthonormal basis changes) gives the passive viewpoint, while the second definition (as those matrices which represent isometries) gives the active viewpoint.

Example 4.7 *The special unitary and orthogonal groups $SU(n)$ and $SO(n)$*

The groups $O(n)$ and $U(n)$ have some very important subgroups, the **special unitary** and **special orthogonal** groups, denoted $SU(n)$ and $SO(n)$, respectively, which are defined as those matrices in $U(n)$ and $O(n)$ that have determinant equal to 1. You will verify below that these are subgroups of $U(n)$ and $O(n)$. These groups are basic in mathematics and (for certain n) fundamental in physics: as we will see, $SO(n)$ is the group of rotations in n dimensions, $SU(2)$ is crucial in the theory of angular momentum in quantum mechanics, and though we will not discuss it here, $SU(3)$ is fundamental in particle physics, especially in the mathematical description of quarks.

Exercise 4.6 Show that $SO(n)$ and $SU(n)$ are subgroups of $O(n)$ and $U(n)$.

Before moving on to a more detailed look at some specific instances of the groups described above, we switch gears for a moment and consider groups that are not subsets of $GL(n, C)$. These groups have a very different flavor than the groups we

have been considering, but are useful in physics nonetheless. We will make more precise the sense in which they differ from the previous examples when we get to Sect. 4.4.

Example 4.8 \mathbb{Z}_2 : *The group with two elements*

Consider the set $\mathbb{Z}_2 \equiv \{+1, -1\} \subset \mathbb{Z}$ with the product being just the usual multiplication of integers. You can easily check that this is a group, in fact an abelian group. Though this group may seem trivial and somewhat abstract, it pops up in a few places in physics, as we will see in Sect. 4.3.

Example 4.9 S_n : *The symmetric group on n letters*

This group does not usually occur explicitly in physics but is intimately tied to permutation symmetry, the physics of identical particles, and much of the mathematics we discussed in Sect. 3.8. The **symmetric group on n letters** (also known as the **permutation group**), denoted S_n , is defined to be the set of all one-to-one and onto maps of the set $\{1, 2, \dots, n\}$ to itself, where the product is just the composition of maps. The maps are known as **permutations**. You should check that any composition of permutations is again a permutation and that permutations are invertible, so that S_n is a group. This verification is simple, and just relies on the fact that permutations are, by definition, one-to-one and onto.

Any permutation σ is specified by the n numbers $\sigma(i)$, $i = 1, \dots, n$, and can conveniently be notated as

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

In such a scheme, the identity in S_3 would just look like

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

while the cyclic permutation σ_1 given by $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 1$ would look like

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

and the transposition σ_2 which switches 1 and 2 and leaves 3 alone would look like

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

How do we take products of permutations? Well, the product $\sigma_1 \cdot \sigma_2$ would take on the following values:

$$(\sigma_1 \cdot \sigma_2)(1) = \sigma_1(\sigma_2(1)) = \sigma_1(2) = 3 \quad (4.17)$$

$$(\sigma_1 \cdot \sigma_2)(2) = \sigma_1(1) = 2 \quad (4.18)$$

$$(\sigma_1 \cdot \sigma_2)(3) = \sigma_1(3) = 1 \quad (4.19)$$

so we have

$$\sigma_1 \cdot \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}. \quad (4.20)$$

You should take the time to inspect (4.20) and understand how to take such a product of permutations without having to write out (4.17)–(4.19).

Though a proper discussion of the applications of S_n to physics must wait until Sect. 4.3, we point out here that if we have a vector space V and consider its n -fold tensor product $\mathcal{T}_n^0(V)$, then S_n acts on product states by

$$\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}.$$

A generic element of $\mathcal{T}_n^0(V)$ will be a sum of such product states, and the action of $\sigma \in S_n$ on these more general states is determined by imposing the linearity condition. In the case of n identical particles in quantum mechanics, where the total Hilbert space is naively the n -fold tensor product $\mathcal{T}_n^0(\mathcal{H})$ of the single-particle Hilbert space \mathcal{H} , this action effectively *interchanges particles*, and we will later restate the symmetrization postulate from Example 3.25 in terms of this action of S_n on $\mathcal{T}_n^0(\mathcal{H})$.

Exercise 4.7 Show that S_n has $n!$ elements.

4.2 The Groups of Classical and Quantum Physics

We are now ready for a detailed look at some of the specific groups which arise in physics.

Example 4.10 $SO(2)$: *Special orthogonal group in two dimensions*

As discussed above, $SO(2)$ is the group of all orthogonal 2×2 matrices with determinant equal to 1. You will check in Exercise 4.8 that $SO(2)$ is abelian and that the general form of an element of $SO(2)$ is

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (4.21)$$

You will recognize that such a matrix represents a counterclockwise rotation of θ radians in the x – y plane. Though we will not discuss $SO(2)$ very much, it serves as a nice warmup for the next example, which is ubiquitous in physics and will be discussed throughout the text.

Exercise 4.8 Consider an arbitrary matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and impose the orthogonality condition, as well as $|A| = 1$. Show that (4.21) is the most general solution to these constraints. Then, verify explicitly that $SO(2)$ is a group (even though we already know it is by Exercise 4.6) by showing that the product of two matrices of the form (4.21) is again a matrix of the form (4.21). This will also show that $SO(2)$ is abelian.

Example 4.11 $SO(3)$: *Special orthogonal group in three dimensions*

This group is of great importance in physics, as it is the group of all rotations in three-dimensional space! For that statement to mean anything, however, we must carefully define what a ‘rotation’ is. One commonly used definition is the following:

Definition A *rotation* in n dimensions is any linear operator R which can be *obtained continuously from the identity*⁵ and takes orthonormal bases to orthonormal bases, i.e. for any orthonormal basis $\{e_i\}_{i=1,\dots,n}$, $\{Re_i\}_{i=1,\dots,n}$ must also be an orthonormal basis.

You will show in Problem 4.1 that this definition is equivalent to saying $R \in SO(n)$.

Given that $SO(3)$ really is the group of three-dimensional rotations, then, can we find a general form for an element of $SO(3)$? As you may know from classical mechanics courses, an arbitrary rotation can be described in terms of the **Euler angles**, which tell us how to rotate a given orthonormal basis into another of the same orientation (or handedness). In classical mechanics texts,⁶ it is shown that this can be achieved by rotating the given axes by an angle ϕ around the original z -axis, then by an angle θ around the *new* x -axis, and finally by an angle ψ around the *new* z -axis. If we take the passive point of view, these three rotations take the form

$$\begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}, \quad \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.22)$$

so multiplying them together gives a general form for $R \in SO(3)$:

$$\begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \psi \sin \theta \\ -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \cos \psi \sin \theta \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix}. \quad (4.23)$$

Another general form for $R \in SO(3)$ is that of a rotation by an arbitrary angle θ about an arbitrary axis $\hat{\mathbf{n}}$; you will see in Sect. 4.6 that this is given by

$$\begin{pmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) - n_z \sin \theta & n_x n_z(1 - \cos \theta) + n_y \sin \theta \\ n_y n_x(1 - \cos \theta) + n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta \\ n_z n_x(1 - \cos \theta) - n_y \sin \theta & n_z n_y(1 - \cos \theta) + n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{pmatrix} \quad (4.24)$$

⁵Meaning that there exists a continuous map $\gamma : [0, 1] \rightarrow GL(n, \mathbb{R})$ such that $\gamma(0) = I$ and $\gamma(1) = R$. In other words, there is a path of *invertible* matrices connecting R to I .

⁶Such as Goldstein [6].

where $\hat{\mathbf{n}} = (n_x, n_y, n_z)$ and the components of $\hat{\mathbf{n}}$ are not all independent since $n_x^2 + n_y^2 + n_z^2 = 1$. This constraint, along with the three components of $\hat{\mathbf{n}}$ and the angle θ , then give us three free parameters with which to describe an arbitrary rotation, just as with the Euler angles. For a nice geometric interpretation of the above matrix, see Problem 4.3.

Example 4.12 $O(3)$: Orthogonal group in 3 dimensions

If $SO(3)$ is the group of all three-dimensional rotations, then what are we to make of $O(3)$, the group of all orthogonal 3×3 matrices without the restriction on the determinant? Well, as we pointed out in Example 3.27, the orthogonality condition actually implies⁷ that $|R| = \pm 1$, so in going from $SO(3)$ to $O(3)$ we are just adding all the orthogonal matrices with $|R| = -1$. These new matrices are sometimes referred to as **improper rotations**, as opposed to the elements with $|R| = 1$ which are known as **proper rotations**. Now, amongst the improper rotations is our old friend the **inversion** transformation, $-I$, which should be familiar from Example 3.27. Any improper rotation can be written as the product of a proper rotation and the inversion transformation, as $R = (-I)(-R)$ (note that if R is an improper rotation, then $-R$ is a proper rotation). Thus, an improper rotation can be thought of as a proper rotation followed⁸ by the inversion transformation. One important feature of $O(3)$ is that its two parts, the proper and improper rotations, are **disconnected**, in the sense that one cannot continuously go from matrices with $|R| = 1$ to matrices with $|R| = -1$ (if one *can* continuously go from one group element to any other, then the group is said to be **connected**. It is disconnected if it is not connected). One can, however, multiply by $-I$ to go between the two components. This is represented schematically in Fig. 4.1. Note that the stipulation in our definition that a rotation must be continuously obtainable from the identity excludes all the improper rotations, as it should.

Example 4.13 $SU(2)$: Special unitary group in two complex dimensions

As mentioned in Example 4.7, $SU(2)$ is the group of all 2×2 complex matrices A which satisfy $|A| = 1$ and

$$A^\dagger = A^{-1}.$$

You can check (see Exercise 4.9 below) that a generic element of $SU(2)$ looks like

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad \alpha, \beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (4.25)$$

⁷This fact can be understood geometrically: since orthogonal matrices preserve distances and angles, they should preserve volumes as well. As we learned in Example 3.27, the determinant measures how volume changes under the action of a linear operator, so any volume preserving operator should have determinant ± 1 . The sign is determined by whether or not the orientation is reversed.

⁸One can actually think of the inversion as following *or* preceding the proper rotation, since $-I$ commutes with all matrices.

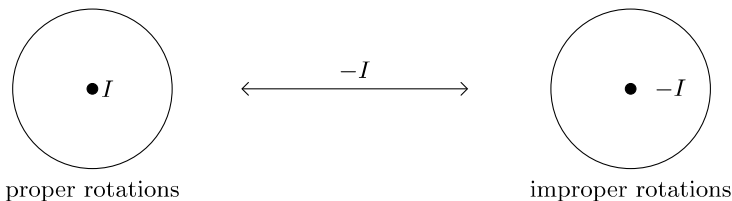


Fig. 4.1 The two components of $O(3)$. The proper rotations are just $SO(3)$. Multiplying by the inversion transformation $-I$ takes one back and forth between the two components

We could also use three real parameters with no conditions rather than two complex parameters with a constraint; one such parametrization is

$$\begin{pmatrix} e^{i(\psi+\phi)/2} \cos \frac{\theta}{2} & i e^{i(\psi-\phi)/2} \sin \frac{\theta}{2} \\ i e^{-i(\psi-\phi)/2} \sin \frac{\theta}{2} & e^{-i(\psi+\phi)/2} \cos \frac{\theta}{2} \end{pmatrix} \quad (4.26)$$

where we have used the same symbols for our parameters as we did for the Euler angles. This is no accident, as there is a close relationship between $SU(2)$ and $SO(3)$, which we will discuss in detail in the next section. This relationship underlies the appearance of $SU(2)$ in quantum mechanics, where rotations are implemented on spin $1/2$ particles by elements of $SU(2)$, and in fact a rotation with Euler angles ϕ , θ and ψ is implemented by the matrix (4.26)!

Exercise 4.9 Consider an arbitrary complex matrix

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

and impose the unit determinant and unitary conditions. Show that (4.25) is the most general solution to these constraints. Then show that any such solution can also be written in the form (4.26).

Example 4.14 $SO(3, 1)_o$: The restricted Lorentz group

The **restricted Lorentz group** $SO(3, 1)_o$ is defined to be the set of all $A \in O(3, 1)$ which satisfy $|A| = 1$ as well as $A_{44} > 1$. You will verify in Problem 4.4 that $SO(3, 1)_o$ is a subgroup of $O(3, 1)$. Where does its definition come from? Well, just as $O(3)$ can be interpreted physically as the set of all orthonormal coordinate transformations, $O(3, 1)$ can be interpreted as the set of all transformations between inertial reference frames. However, we often are interested in restricting those changes of reference frame to those which preserve the orientation of time and space, which is what the additional conditions $|A| = 1$ and $A_{44} > 1$ do. The condition $A_{44} > 0$ means that A does not reverse the direction of time (so that clocks in the new coordinates are not running backwards) and this, together with $|A| = 1$, implies that A does not reverse the orientation of the space axes. Such transformations are known as **restricted Lorentz transformations**.

The most familiar such transformation is probably

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -\beta\gamma \\ 0 & 0 & -\beta\gamma & \gamma \end{pmatrix}, \quad -1 < \beta < 1, \quad \gamma \equiv \frac{1}{\sqrt{1-\beta^2}}. \quad (4.27)$$

This is interpreted passively⁹ as a coordinate transformation to a new reference frame that is unrotated relative to the old frame but is moving uniformly along the z -axis with relative velocity β .¹⁰ Such a transformation is often referred to as a **boost** along the z -axis, and is also sometimes written as

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh u & -\sinh u \\ 0 & 0 & -\sinh u & \cosh u \end{pmatrix}, \quad u \in \mathbb{R} \quad (4.28)$$

where u is a quantity known as the **rapidity** and is related to β by $\tanh u = \beta$. (You should check that the above matrices really are elements of $SO(3, 1)_o$.) We could also boost along any other spatial direction; if the relative velocity vector is $\boldsymbol{\beta} \equiv (\beta_x, \beta_y, \beta_z)$, then the corresponding matrix should be obtainable from (4.27) by an orthogonal similarity transformation that takes $\hat{\mathbf{z}}$ into $\hat{\boldsymbol{\beta}}$. You will show in Exercise 4.10 below that this yields

$$L = \begin{pmatrix} \frac{\beta_x^2(\gamma-1)}{\beta^2} + 1 & \frac{\beta_x\beta_y(\gamma-1)}{\beta^2} & \frac{\beta_x\beta_z(\gamma-1)}{\beta^2} & -\beta_x\gamma \\ \frac{\beta_y\beta_x(\gamma-1)}{\beta^2} & \frac{\beta_y^2(\gamma-1)}{\beta^2} + 1 & \frac{\beta_y\beta_z(\gamma-1)}{\beta^2} & -\beta_y\gamma \\ \frac{\beta_z\beta_x(\gamma-1)}{\beta^2} & \frac{\beta_z\beta_y(\gamma-1)}{\beta^2} & \frac{\beta_z^2(\gamma-1)}{\beta^2} + 1 & -\beta_z\gamma \\ -\beta_x\gamma & -\beta_y\gamma & -\beta_z\gamma & \gamma \end{pmatrix}. \quad (4.29)$$

If we generalize the relation between u and $\boldsymbol{\beta}$ to three dimensions as

$$\boldsymbol{\beta} = \frac{\tanh u}{u} \mathbf{u}, \quad u = |\mathbf{u}| \quad (4.30)$$

⁹It is worth noting that, in contrast to rotations, Lorentz transformations are pretty much *always* interpreted passively. A vector in \mathbb{R}^4 is considered an event, and it does not make much sense to start moving that event around in spacetime (the active interpretation), though it does make sense to ask what a different observer's coordinates for that particular event would be (passive interpretation).

¹⁰Note that β is measured in units of the speed of light, hence the restriction $-1 < \beta < 1$.

then you can check that this arbitrary boost can also be written as

$$L = \begin{pmatrix} \frac{u_x^2(\cosh u - 1)}{u^2} + 1 & \frac{u_x u_y(\cosh u - 1)}{u^2} & \frac{u_x u_z(\cosh u - 1)}{u^2} & -\frac{u_x}{u} \sinh u \\ \frac{u_y u_x(\cosh u - 1)}{u^2} & \frac{u_y^2(\cosh u - 1)}{u^2} + 1 & \frac{u_y u_z(\cosh u - 1)}{u^2} & -\frac{u_y}{u} \sinh u \\ \frac{u_z u_x(\cosh u - 1)}{u^2} & \frac{u_z u_y(\cosh u - 1)}{u^2} & \frac{u_z^2(\cosh u - 1)}{u^2} + 1 & -\frac{u_z}{u} \sinh u \\ -\frac{u_x}{u} \sinh u & -\frac{u_y}{u} \sinh u & -\frac{u_z}{u} \sinh u & \cosh u \end{pmatrix}. \quad (4.31)$$

Now we know what boosts look like, but how about an arbitrary restricted Lorentz transformation? Well, the nice thing about $SO(3, 1)_o$ is that any element A can be decomposed as $A = LR'$, where

$$R' = \begin{pmatrix} R & \\ & 1 \end{pmatrix}, \quad R \in SO(3) \quad (4.32)$$

and L is of the form (4.29). This is the usual decomposition of an arbitrary restricted Lorentz transformation into a rotation and a boost, which you will perform in Problem 4.5. Note that L has three arbitrary parameters, so that our arbitrary restricted Lorentz transformation LR' has six parameters total.

Exercise 4.10 Construct an orthogonal matrix R which implements an orthonormal change of basis from the standard basis $\{\hat{x}, \hat{y}, \hat{z}\}$ to one of the form $\{\hat{r}_1, \hat{r}_2, \hat{\beta}\}$ where the \hat{r}_i are any two vectors mutually orthonormal with $\hat{\beta}$ and each other. Embed R in $SO(3, 1)_o$ as in (4.32) and use this to obtain (4.29) by performing a similarity transformation on (4.27). Parts of Problem 3.1 may be useful here.

Exercise 4.11 Check that L in (4.29) really does represent a boost of velocity β as follows: Use L as a passive transformation to obtain new coordinates (x', y', z', t') from the old ones by

$$\begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} = L \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}.$$

Show that the spatial origin of the unprimed frame, defined by $x = y = z = 0$, moves with velocity $-\beta$ in the primed coordinate system, which tells us that the primed coordinate system moves with velocity $+\beta$ with respect to the unprimed system.

Example 4.15 $O(3, 1)$: The extended Lorentz group

In the previous example we restricted our changes of inertial reference frame to those which preserved the orientation of space and time. This is sufficient in classical mechanics, but in quantum mechanics we are often interested in the effects of space and time inversion on the various Hilbert spaces we are working with. If we add spatial inversion, also called **parity** and represented by the matrix

$$P = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4.33)$$

as well as time-reversal, represented by

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (4.34)$$

to the restricted Lorentz group, we actually recover $O(3, 1)$, which is thus known as the **improper** or **extended** Lorentz group. You should verify that $P, T \in O(3, 1)$, but $P, T \notin SO(3, 1)_o$. In fact, $|P| = |T| = -1$, which is no accident; as in the case of the orthogonal group, the defining equation (4.14) restricts the determinant, and in fact implies that $|A| = \pm 1$. In this case, however, the group has four connected components instead of two! Obviously those matrices with $|A| = 1$ must be disconnected from those with $|A| = -1$, but those which reverse the orientation of the space axes must also be disconnected from those which do not, and those which reverse the orientation of time must be disconnected from those which do not. This is represented schematically in Fig. 4.2. Note that, as in the case of $O(3)$, multiplication by the transformations P and T take us to and from the various different components.

Example 4.16 $SL(2, \mathbb{C})$: *Special linear group in two complex dimensions*

This group cannot be viewed as a group of isometries, but it is important in physics nonetheless. $SL(2, \mathbb{C})$ is defined to be the set of all 2×2 complex matrices A with $|A| = 1$. By now it should be obvious that this set is a group. The general form of $A \in SL(2, \mathbb{C})$ is

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{C}, \quad ad - bc = 1.$$

The unit determinant constraint means that A is determined by three complex parameters or six real parameters, just as for $SO(3, 1)_o$. This is no coincidence; in fact, $SL(2, \mathbb{C})$ bears the same relationship to $SO(3, 1)_o$ as $SU(2)$ bears to $SO(3)$, in that $SL(2, \mathbb{C})$ implements restricted Lorentz transformations on spin $1/2$ particles! This will be discussed in the next section. You will also show later¹¹ that a boost with rapidity \mathbf{u} is implemented by an $SL(2, \mathbb{C})$ matrix of the form

$$\tilde{L} = \begin{pmatrix} \cosh \frac{u}{2} + \frac{u_z}{u} \sinh \frac{u}{2} & -\frac{1}{u}(u_x - iu_y) \sinh \frac{u}{2} \\ -\frac{1}{u}(u_x + iu_y) \sinh \frac{u}{2} & \cosh \frac{u}{2} - \frac{u_z}{u} \sinh \frac{u}{2} \end{pmatrix}, \quad \mathbf{u} \in \mathbb{R}^3. \quad (4.35)$$

It can also be shown,¹² just as for $SO(3, 1)_o$, that any $A \in SL(2, \mathbb{C})$ can be decomposed as $A = \tilde{L}\tilde{R}$, where $\tilde{R} \in SU(2)$ and \tilde{L} is as above. This, together with the facts that an arbitrary rotation can be implemented by $\tilde{R} \in SU(2)$ parametrized as in (4.26), yields the general form $\tilde{L}\tilde{R}$ for an element of $SL(2, \mathbb{C})$ in terms of the same parameters we used for $SO(3, 1)_o$.

¹¹See Problem 4.8.

¹²See Problem 4.6.

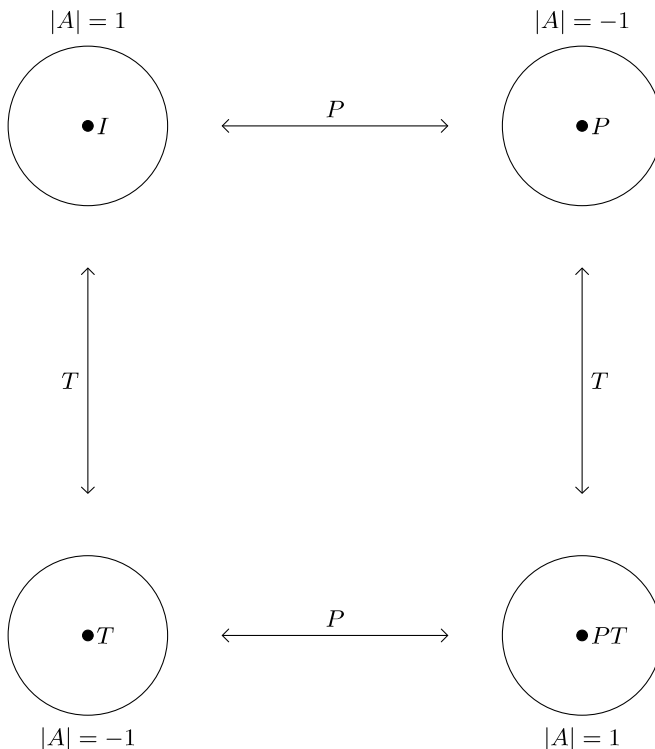


Fig. 4.2 The four components of $O(3, 1)$. The proper Lorentz transformations in the upper-left corner are just $SO(3, 1)_o$. Note that the transformations in the lower-right hand corner change both the orientation of the space axes as well as time, and so must be disconnected from $SO(3, 1)_o$ even though they have $|A| = 1$. Also note that multiplying by the parity and time-reversal operators P and T take one back and forth between the various components

4.3 Homomorphism and Isomorphism

In the last section we claimed that there is a close relationship between $SU(2)$ and $SO(3)$, as well as between $SL(2, \mathbb{C})$ and $SO(3, 1)_o$. We now make this relationship precise, and show that a similar relationship exists between S_n and \mathbb{Z}_2 . We will also define what it means for two groups to be ‘the same’, which will then tie into our somewhat abstract discussion of \mathbb{Z}_2 in the last section.

Given two groups G and H , a **homomorphism** from G to H is a map $\Phi : G \rightarrow H$ such that

$$\boxed{\Phi(g_1 g_2) = \Phi(g_1) \Phi(g_2) \quad \forall g_1, g_2 \in G.} \quad (4.36)$$

Note that the product in the left hand side of (4.36) takes place in G , whereas the product on the right hand side takes place in H . A homomorphism should be thought of as a map from one group to another which preserves the multiplicative structure. Note that Φ need not be one-to-one or onto; if it is onto then Φ is said to be a

homomorphism onto H , and if in addition it is one-to-one then we say Φ is an **isomorphism**. If Φ is an isomorphism then it is invertible and thus sets up a one-to-one correspondence which preserves the group structure, so we regard G and H as ‘the same’ group, just with different labels for the elements. When two groups G and H are isomorphic we write $G \simeq H$.

Exercise 4.12 Let $\Phi : G \rightarrow H$ be a homomorphism, and let e be the identity in G and e' the identity in H . Show that

$$\begin{aligned}\Phi(e) &= e' \\ \Phi(g^{-1}) &= \Phi(g)^{-1} \quad \forall g \in G.\end{aligned}$$

Example 4.17 *Isometries and the orthogonal, unitary, and Lorentz groups*

A nice example of a group isomorphism is when we have an n -dimensional vector space V (over some scalars C) and a basis \mathcal{B} , hence a map

$$\begin{aligned}GL(V) &\rightarrow GL(n, C) \\ T &\mapsto [T]_{\mathcal{B}}.\end{aligned}$$

It is easily checked that this map is one-to-one and onto. Furthermore, it is a homomorphism since $[TU] = [T][U]$, a fact you proved in Exercise 2.8. Thus this map is an isomorphism and $GL(V) \simeq GL(n, C)$. If V has a non-degenerate Hermitian form $(\cdot | \cdot)$ then we can restrict this map to $\text{Isom}(V) \subset GL(V)$, which yields

$$\text{Isom}(V) \simeq O(n)$$

when V is real and $(\cdot | \cdot)$ is positive-definite,

$$\text{Isom}(V) \simeq U(n)$$

when V is complex and $(\cdot | \cdot)$ is positive-definite, and

$$\text{Isom}(V) \simeq O(n-1, 1)$$

when V is real and $(\cdot | \cdot)$ is a Minkowski metric. These isomorphisms were implicit in the discussion of Examples 4.3–4.6, where we identified the operators in $\text{Isom}(V)$ with their matrix representations in the corresponding matrix group. \square

Example 4.18 *Linear maps as homomorphisms*

A **linear map** from a vector space V to a vector space W is a map $\Phi : V \rightarrow W$ that satisfies the usual linearity condition

$$\Phi(cv_1 + v_2) = c\Phi(v_1) + \Phi(v_2). \quad (4.37)$$

(A linear *operator* is then just the special case in which $V = W$.) In particular we have $\Phi(v_1 + v_2) = \Phi(v_1) + \Phi(v_2)$, which just says that Φ is a homomorphism between the additive groups V and W ! (cf. Example 4.2). If Φ is one-to-one and

onto then it is an isomorphism, and in particular we refer to it as a **vector space isomorphism**.

The notions of linear map and vector space isomorphism are basic ones, and could have been introduced much earlier (as they are in standard linear algebra texts). Because of our specific goals in this book, however, we have not needed them yet. These objects will start to play a role soon, though, and will recur throughout the rest of the book.

Exercise 4.13 Use an argument similar to that of Exercise 2.6 to prove that a linear map $\phi : V \rightarrow W$ is an isomorphism if and only if $\dim V = \dim W$ and ϕ satisfies

$$\phi(v) = 0 \implies v = 0.$$

Exercise 4.14 Before moving on to more complicated examples, let us get some practice by acquainting ourselves with a few more basic homomorphisms. First, show that the map

$$\begin{aligned} \exp : \mathbb{R} &\rightarrow \mathbb{R}^* \\ x &\mapsto e^x, \end{aligned}$$

from the additive group of real numbers to the multiplicative group of nonzero real numbers, is a homomorphism. Is it an isomorphism? Why or why not? Repeat the analysis for $\exp : \mathbb{C} \rightarrow \mathbb{C}^*$. Also, show that the map

$$\begin{aligned} \det : GL(n, \mathbb{C}) &\rightarrow \mathbb{C}^* \\ A &\mapsto \det A \end{aligned}$$

is a homomorphism for both $\mathbb{C} = \mathbb{R}$ and $\mathbb{C} = \mathbb{C}$. Is it an isomorphism in either case? Would you expect it to be?

Exercise 4.15 Recall that $U(1)$ is the group of 1×1 unitary matrices. Show that this is just the set of complex numbers z with $|z| = 1$, and that $U(1)$ is isomorphic to $SO(2)$.

Suppose Φ is a homomorphism but not one-to-one. Is there a way to quantify how far it is from being one-to-one? Define the **kernel** of Φ to be the set $K \equiv \{g \in G \mid \Phi(g) = e'\}$, where e' is the identity in H . In other words, K is the set of all elements of G that get sent to e' under Φ . If Φ is one-to-one, then $K = \{e\}$, since there can be only one element in G that maps to e' and by Exercise 4.12 that is e . If Φ is not one-to-one, then the size of K tells us how far it is from being so. Also, if we have $\Phi(g_1) = \Phi(g_2) = h \in H$, then

$$\Phi(g_1 g_2^{-1}) = \Phi(g_1) \Phi(g_2)^{-1} = h h^{-1} = e$$

so $g_1 g_2^{-1}$ is in the kernel of Φ , i.e. $g_1 g_2^{-1} = k \in K$. Multiplying this on the right by g_2 then gives $g_1 = k g_2$, so we see that any two elements of G that give the same element of H under Φ are related by left multiplication by an element of K . Conversely, if we are given $g \in G$ and $\Phi(g) = h$, then for all $k \in K$,

$$\Phi(kg) = \Phi(k) \Phi(g) = e' \Phi(g) = \Phi(g) = h.$$

Thus if we define $Kg \equiv \{kg \mid k \in K\}$, then Kg are *precisely* those elements (no more, and no less) of G which get sent to h . Thus the size of K tells us how far Φ

is from being one-to-one, and the elements of K tell us exactly which elements of G will map to a specific element $h \in H$.

Homomorphisms and isomorphisms are ubiquitous in mathematics and occur frequently in physics, as we will see in the examples below.

Exercise 4.16 Show that the kernel K of any homomorphism $\Phi : G \rightarrow H$ is a subgroup of G . Then determine the kernels of the maps \exp and \det of Exercise 4.14.

Exercise 4.17 Suppose $\Phi : V \rightarrow W$ is a linear map between vector spaces, hence a homomorphism between abelian groups. Conclude from the previous exercise that the kernel K of Φ is a subspace of V , also known as the **null space** of Φ . The dimension of K is known as the **nullity** of K . Also show that the range of Φ is a subspace of W , whose dimension is known as the **rank** of Φ . Finally, prove the **rank-nullity** theorem of linear algebra, which states that

$$\text{rank}(\Phi) + \text{nullity}(\Phi) = \dim V. \quad (4.38)$$

(Hint: Take a basis $\{e_1, \dots, e_k\}$ for K and complete it to a basis $\{e_1, \dots, e_n\}$ for V , where $n = \dim V$. Then show that $\{\Phi(e_{k+1}), \dots, \Phi(e_n)\}$ is a basis for the range of Φ .)

Example 4.19 $SU(2)$ and $SO(3)$

In most physics textbooks the relationship between $SO(3)$ and $SU(2)$ is described in terms of the ‘infinitesimal generators’ of these groups. We will discuss infinitesimal transformations in the next section and make contact with the standard physics presentation then; here we present the relationship in terms of a group homomorphism $\rho : SU(2) \rightarrow SO(3)$, defined as follows: consider the vector space (check!) of all 2×2 traceless anti-Hermitian matrices, denoted as $\mathfrak{su}(2)$ (for reasons we will explain later). You can check that an arbitrary element $X \in \mathfrak{su}(2)$ can be written as

$$X = \frac{1}{2} \begin{pmatrix} -iz & -y - ix \\ y - ix & iz \end{pmatrix}, \quad x, y, z \in \mathbb{R}. \quad (4.39)$$

If we take as basis vectors

$$\begin{aligned} S_x &\equiv -\frac{i}{2}\sigma_x = \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ S_y &\equiv -\frac{i}{2}\sigma_y = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ S_z &\equiv -\frac{i}{2}\sigma_z = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \end{aligned}$$

then we have

$$X = xS_x + yS_y + zS_z$$

so the column vector corresponding to X in the basis $\mathcal{B} = \{S_x, S_y, S_z\}$ is

$$[X] = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Note that

$$\det X = \frac{1}{4}(x^2 + y^2 + z^2) = \frac{1}{4}\| [X] \|^2$$

so the determinant of $X \in \mathfrak{su}(2)$ is proportional to the norm squared of $[X] \in \mathbb{R}^3$ with the usual Euclidean metric. Now, you will check below that $A \in SU(2)$ acts on $X \in \mathfrak{su}(2)$ by the map $X \mapsto AXA^\dagger$, and that this map is linear. Thus, this map is a linear operator on $\mathfrak{su}(2)$, and can be represented in the basis \mathcal{B} by a 3×3 matrix which we will call $\rho(A)$, so that $[AXA^\dagger] = \rho(A)[X]$ where $\rho(A)$ acts on $[X]$ by the usual matrix multiplication. Furthermore,

$$\| \rho(A)[X] \|^2 = \| [AXA^\dagger] \|^2 = 4 \det(AXA^\dagger) = 4 \det X = \| [X] \|^2 \quad (4.40)$$

so $\rho(A)$ preserves the norm of X . This implies (see Exercise 4.19 below) that $\rho(A) \in O(3)$, and one can in fact show¹³ that $\det \rho(A) = 1$, so that $\rho(A) \in SO(3)$. Thus we may construct a map

$$\begin{aligned} \rho : SU(2) &\rightarrow SO(3) \\ A &\mapsto \rho(A). \end{aligned}$$

Furthermore, ρ is a homomorphism, since

$$\begin{aligned} \rho(AB)[X] &= [(AB)X(AB)^\dagger] = [ABXB^\dagger A^\dagger] = \rho(A)[BXB^\dagger] \\ &= \rho(A)\rho(B)[X] \end{aligned} \quad (4.41)$$

and hence $\rho(AB) = \rho(A)\rho(B)$. Is ρ an isomorphism? One can show¹⁴ that ρ is onto but not one-to-one, and in fact has kernel $K = \{I, -I\}$. From the discussion preceding this example, we then know that $\rho(A) = \rho(-A) \forall A \in SU(2)$ (this fact is also clear from the definition of ρ), so for every rotation $R \in SO(3)$ there correspond exactly two matrices in $SU(2)$ which map to R under ρ . Thus, when trying to implement a rotation R on a spin $1/2$ particle we have two choices for the $SU(2)$ matrix we use, and it is sometimes said that the map ρ^{-1} from $SO(3)$ to $SU(2)$ is *double-valued*. In mathematical terms one does not usually speak of functions with multiple-values, though, so instead we say that $SU(2)$ is the **double cover** of $SO(3)$, since the map ρ is onto ('cover') and two-to-one ('double').

¹³See Problem 4.7 of this chapter, or consider the following rough (but correct) argument: $\rho : SU(2) \rightarrow O(3)$ as defined above is a continuous map, and so the composition

$$\begin{aligned} \det \circ \rho : SU(2) &\rightarrow \mathbb{R} \\ A &\mapsto \det(\rho(A)) \end{aligned}$$

is also continuous. Since $SU(2)$ is connected any continuous function must itself vary continuously, so $\det \circ \rho$ cannot jump between 1 and -1 , which are its only possible values. Since $\det(\rho(I)) = 1$, we can then conclude that $\det(\rho(A)) = 1 \forall A \in SU(2)$.

¹⁴See Problem 4.7 again.

Exercise 4.18 Let $A \in SU(2)$, $X \in \mathfrak{su}(2)$. Show that $AXA^\dagger \in \mathfrak{su}(2)$ and note that $A(X + Y)A^\dagger = AXA^\dagger + AYA^\dagger$, so that the map $X \rightarrow AXA^\dagger$ really is a linear operator on $\mathfrak{su}(2)$.

Exercise 4.19 Let V be a real vector space with a metric g and let $R \in \mathcal{L}(V)$ preserve norms on V , i.e. $g(Rv, Rv) = g(v, v) \forall v \in V$. Show that this implies that

$$g(Rv, Rw) = g(v, w) \quad \forall v, w \in V,$$

i.e. that R is an isometry. Hint: consider $g(v + w, v + w)$ and use the bilinearity of g .

Example 4.20 $SL(2, \mathbb{C})$ and $SO(3, 1)_o$

Just as there is a two-to-one homomorphism from $SU(2)$ to $SO(3)$, there is a two-to-one homomorphism from $SL(2, \mathbb{C})$ to $SO(3, 1)_o$ which is defined similarly. Consider the vector space $H_2(\mathbb{C})$ of 2×2 Hermitian matrices. As we saw in Example 2.9, this is a four-dimensional vector space with basis $\mathcal{B} = \{\sigma_x, \sigma_y, \sigma_z, I\}$, and an arbitrary element $X \in H_2(\mathbb{C})$ can be written as

$$X = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} = x\sigma_x + y\sigma_y + z\sigma_z + tI \quad (4.42)$$

so that

$$[X] = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}.$$

Now, $SL(2, \mathbb{C})$ acts on $H_2(\mathbb{C})$ in the same way that $SU(2)$ acts on $\mathfrak{su}(2)$: by sending $X \rightarrow AXA^\dagger$ where $A \in SL(2, \mathbb{C})$. You can again check that this is actually a linear map from $H_2(\mathbb{C})$ to itself, and hence can be represented in components by a matrix which we will again call $\rho(A)$. You can also check that

$$\det X = t^2 - x^2 - y^2 - z^2 = -\eta([X], [X])$$

so the determinant of $X \in H_2(\mathbb{C})$ gives minus the norm squared of $[X]$ in the *Minkowski* metric on \mathbb{R}^4 . As before, the action of $\rho(A)$ on $[X]$ preserves this norm (by a calculation identical to (4.40)), and you will show in Problem 4.8 that $\det \rho(A) = 1$ and $\rho(A)_{44} > 1$, so $\rho(A) \in SO(3, 1)_o$. Thus we can again construct a map

$$\begin{aligned} \rho : SL(2, \mathbb{C}) &\rightarrow SO(3, 1)_o \\ A &\mapsto \rho(A) \end{aligned}$$

and it can be shown¹⁵ that ρ is onto. Furthermore, ρ is a homomorphism, by a calculation identical to (4.41). The kernel of ρ is again $K = \{I, -I\}$ so $SL(2, \mathbb{C})$ is a double-cover of $SO(3, 1)_o$.

¹⁵See Problem 4.8 again.

Exercise 4.20 The elements of $SL(2, \mathbb{C})$ that correspond to rotations should fix timelike vectors, i.e. leave them unchanged. Identify the elements of $H_2(\mathbb{C})$ that correspond to timelike vectors, and show that it is precisely the subgroup $SU(2) \subset SL(2, \mathbb{C})$ which leaves them unchanged, as expected.

Example 4.21 \mathbb{Z}_2 , parity and time-reversal

Consider the set $\{I, P\} \subset O(3, 1)$. This is an abelian two-element group with $P^2 = I$, and so looks just like \mathbb{Z}_2 . In fact, if we define a map $\Phi : \{I, P\} \rightarrow \mathbb{Z}_2$ by

$$\begin{aligned}\Phi(I) &= 1 \\ \Phi(P) &= -1\end{aligned}$$

then Φ is a homomorphism since

$$\Phi(P \cdot P) = \Phi(I) = 1 = (-1)^2 = \Phi(P)\Phi(P).$$

Φ is also clearly one-to-one and onto, so Φ is in fact an isomorphism! We could also consider the two-element group $\{I, T\} \in O(3, 1)$; since $T^2 = I$, we could define a similar isomorphism from $\{I, T\}$ to \mathbb{Z}_2 . Thus,

$$\mathbb{Z}_2 \simeq \{I, P\} \simeq \{I, T\}.$$

In fact, you will show below that *all* two-element groups are isomorphic. That is why we chose to present \mathbb{Z}_2 somewhat abstractly; there are many groups in physics that are isomorphic to \mathbb{Z}_2 , so it makes sense to use an abstract formulation so that we can talk about all of them at the same time without having to think in terms of a particular representation like $\{I, P\}$ or $\{I, T\}$. We will see the advantage of this in the next example.

Exercise 4.21 Show that any group G with only two elements e and g must be isomorphic to \mathbb{Z}_2 . To do this you must define a map $\Phi : G \rightarrow \mathbb{Z}_2$ which is one-to-one, onto, and which satisfies (4.36). Note that S_2 , the symmetric group on two letters, has only two elements. What is the element that corresponds to $-1 \in \mathbb{Z}_2$?

Example 4.22 S_n , \mathbb{Z}_2 , and the sgn homomorphism

In Sect. 3.8 we discussed rearrangements, transpositions, and the evenness or oddness of a rearrangement in terms of the number of transpositions needed to obtain it. We are now in a position to make this much more precise, which will facilitate neater descriptions of the ϵ tensor, the determinant of a matrix, and the symmetrization postulate.

We formally define a **transposition** in S_n to be any permutation σ which switches two numbers i and j and leaves all the others alone. You can check that σ_2 and $\sigma_1 \cdot \sigma_2$ from Example 4.9 are both transpositions. It is a fact that *any* permutation can be written (non-uniquely) as a product of transpositions, though we will not prove this here.¹⁶ As a simple example, though, you can check that

¹⁶See Herstein [9] for a proof and nice discussion.

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

is a decomposition of σ_1 into transpositions.

Though the decomposition of a given permutation is far from unique (for instance, the identity can be decomposed as a product of any transposition σ and its inverse), the *evenness* or *oddness* of the number of transpositions in a decomposition is invariant. For instance, even though we could write the identity as

$$\begin{aligned} e &= e \\ e &= \sigma_1 \sigma_1^{-1} \\ e &= \sigma_1 \sigma_1^{-1} \sigma_2 \sigma_2^{-1} \end{aligned}$$

and so on for any transpositions $\sigma_1, \sigma_2, \dots$, every decomposition will consist of an *even* number of transpositions. Likewise, any decomposition of the transposition $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ will consist of an odd number of transpositions. A general proof of this fact is relegated to Problem 4.9. What this allows us to do, though, is define a homomorphism $\text{sgn}: S_n \rightarrow \mathbb{Z}_2$ by

$$\text{sgn}(\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ consists of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ consists of an odd number of transpositions.} \end{cases}$$

You should be able to verify with a moment's thought that

$$\text{sgn}(\sigma_1 \sigma_2) = \text{sgn}(\sigma_1) \text{sgn}(\sigma_2),$$

so that sgn is actually a homomorphism. If $\text{sgn}(\sigma) = +1$ we say σ is *even*, and if $\text{sgn}(\sigma) = -1$ we say σ is *odd*.

With the sgn homomorphism in hand we can now tidy up several definitions from Sect. 3.8. First of all, we can now define the wedge product of r dual vectors $f_i, i = 1, \dots, r$ to be

$$f_1 \wedge \cdots \wedge f_r \equiv \sum_{\sigma \in S_r} \text{sgn}(\sigma) f_{\sigma(1)} \otimes f_{\sigma(2)} \otimes \cdots \otimes f_{\sigma(r)}. \quad (4.43)$$

You should compare this with the earlier definition and convince yourself that the two definitions are equivalent. Also, from our earlier definition of the ϵ tensor it should be clear that the nonzero components of ϵ are

$$\epsilon_{i_1 \dots i_n} = \text{sgn}(\sigma) \quad \text{where } \sigma = \begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix} \quad (4.44)$$

and so the definition of the determinant, (3.85), becomes

$$|A| = \sum_{\sigma \in S_n} \text{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}. \quad (4.45)$$

Finally, at the end of Example 4.9 we described how S_n acts on an n -fold tensor product $\mathcal{T}_n^0(V)$ by

$$\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}, \quad (4.46)$$

where the action of σ on a generic element on $\mathcal{T}_n^0(V)$ (which would be a sum of product states like the above) is determined by linearity. If we have a totally symmetric tensor $T = T^{i_1 \dots i_n} e_{i_1} \otimes \dots \otimes e_{i_n} \in S^n(V)$, we then have

$$\begin{aligned} \sigma(T) &= T^{i_1 \dots i_n} e_{\sigma(i_1)} \otimes \dots \otimes e_{\sigma(i_n)} \\ &= T^{\sigma^{-1}(j_1) \dots \sigma^{-1}(j_n)} e_{j_1} \otimes \dots \otimes e_{j_n} \\ &\quad \text{where we have relabeled indices using } j_k \equiv \sigma(i_k) \\ &= T^{j_1 \dots j_n} e_{j_1} \otimes \dots \otimes e_{j_n} \quad \text{by total symmetry of } T^{j_1 \dots j_n} \\ &= T \end{aligned}$$

so all elements of $S^n(V)$ are fixed by the action of S_n . If we now consider a totally antisymmetric tensor $T = T^{i_1 \dots i_n} e_{i_1} \otimes \dots \otimes e_{i_n} \in \Lambda^n(V)$, then the action of $\sigma \in S_n$ on it is given by

$$\begin{aligned} \sigma(T) &= T^{i_1 \dots i_n} e_{\sigma(i_1)} \otimes \dots \otimes e_{\sigma(i_n)} \\ &= T^{\sigma^{-1}(j_1) \dots \sigma^{-1}(j_n)} e_{j_1} \otimes \dots \otimes e_{j_n} \\ &= \text{sgn}(\sigma) T^{j_1 \dots j_n} e_{j_1} \otimes \dots \otimes e_{j_n} \quad \text{by antisymmetry of } T^{j_1 \dots j_n} \\ &= \text{sgn}(\sigma) T \end{aligned}$$

so if σ is odd then T changes sign under it, and if σ is even then T is invariant. Thus, we can restate the symmetrization postulate as follows: any state of an n -particle system is either *invariant* under a permutation of the particles (in which case the particles are known as **bosons**), or *changes sign* depending on whether the permutation is even or odd (in which case the particles are known as **fermions**).

4.4 From Lie Groups to Lie Algebras

You may have noticed that the examples of groups we met in the last two sections had a couple of different flavors: there were the matrix groups like $SU(2)$ and $SO(3)$ which were parametrizable by a certain number of real parameters, and then there were the ‘discrete’ groups like \mathbb{Z}_2 and the symmetric groups S_n that had a finite number of elements and were described by discrete labels rather than continuous parameters. The first type of group forms an extremely important subclass of groups known as **Lie Groups**, named after the Norwegian mathematician Sophus Lie who was among the first to study them systematically in the late 1800s. Besides their ubiquity in math and physics, Lie groups are important because their continuous nature means that we can study group elements that are ‘infinitely close’ to the identity; these are known to physicists as the ‘infinitesimal transformations’ or ‘generators’ of the group, and to mathematicians as the **Lie algebra** of the group. As we make this notion precise, we will see that Lie algebras are vector spaces and as such are sometimes simpler to understand than the ‘finite’ group elements. Also, the ‘generators’ in some Lie algebras are taken in quantum mechanics to represent

certain physical observables, and in fact almost all observables can be built out of elements of certain Lie algebras. We will see that many familiar objects and structures in physics can be understood in terms of Lie algebras.

Before we can study Lie algebras, however, we should make precise what we mean by a Lie group. Here we run into a snag, because the proper and most general definition requires machinery well outside the scope of this text.¹⁷ We do wish to be precise, though, so we follow Hall [8] and use a restricted definition. This definition does not really capture the essence of what a Lie Group is, but will get the job done and allow us to discuss Lie algebras without having to wave our hands.

That said, we define a **matrix Lie group** to be a subgroup $G \subset GL(n, \mathbb{C})$ which is **closed**, in the following sense: for any sequence of matrices $A_n \in G$ which converges to a limit matrix A , either $A \in G$ or $A \notin GL(n, \mathbb{C})$. This says that a limit of matrices in G must either itself be in G , or otherwise be noninvertible. As remarked above, this definition is technical and does not provide much insight into what a Lie group really is, but it will provide the necessary hypotheses in proving the essential properties of Lie algebras.

Let us now prove that some of the groups we have encountered above are indeed matrix Lie groups. We will verify this explicitly for one class of groups, the orthogonal groups, and leave the rest as problems for you. The orthogonal group $O(n)$ is defined by the equation $R^{-1} = R^T$, or $R^T R = I$. Let us consider the function from $GL(n, \mathbb{R})$ to itself defined by $f(A) = A^T A$. Each entry of the matrix $f(A)$ is easily seen to be a continuous function of the entries of A , so f is continuous. Consider now a sequence R_i in $O(n)$ that converges to some limit matrix R . We then have

$$\begin{aligned} f(R) &= f\left(\lim_{i \rightarrow \infty} R_i\right) \\ &= \lim_{i \rightarrow \infty} f(R_i) \quad \text{since } f \text{ is continuous} \\ &= \lim_{i \rightarrow \infty} I \\ &= I \end{aligned}$$

so $R \in O(n)$. Thus $O(n)$ is a matrix Lie group. The unitary and Lorentz groups, as well as their cousins with unit determinant, are similarly defined by continuous functions, and can analogously be shown to be matrix Lie groups. For an example of a subgroup of $GL(n, \mathbb{C})$ which is *not* closed, hence not a matrix Lie group, see Problem 4.10.

We remarked earlier, though, that the above definition does not really capture the essence of what a Lie group is. What is that essence? As mentioned before, one should think of Lie groups as groups which can be *parametrized* in terms of a certain number of real variables. This number is known as the **dimension** of the Lie group, and we will see that this number is also the usual (vector space) dimension of its corresponding Lie algebra. Since a Lie group is parametrizable, we can think of it as

¹⁷The necessary machinery being the theory of differentiable manifolds; in this context, a Lie group is essentially a group that is also a differentiable manifold. See Schutz [15] or Frankel [4] for very readable introductions for physicists, and Warner [18] for a systematic but terse account.

a kind of multi-dimensional space (like the surface of a sphere) that also has a group structure. Unlike the surface of a sphere, though, a Lie group has a distinguished point, the identity e , and as we mentioned above, studying transformations ‘close to’ the identity will lead us to Lie algebras.

For the sake of completeness, we should point out here that there are Lie groups out there which are *not* matrix Lie groups, i.e. which cannot be described as a subset of $GL(n, \mathbb{C})$ for some n . Their relevance for basic physics has not been established, however, so we do not consider them here.¹⁸

Now that we have a better sense of what Lie groups are, we would like to zoom in to Lie algebras by considering group elements that are ‘close’ to the identity. For concreteness consider the rotation group $SO(3)$. An arbitrary rotation about the z axis looks like

$$R_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If we take our rotation angle to be $\epsilon \ll 1$, we can approximate $R_z(\epsilon)$ by expanding to first order, which yields

$$R_z(\epsilon) \approx \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I + \epsilon L_z \quad (4.47)$$

where you may recognize

$$L_z \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

from (2.16), up to a factor of i . The appearance here of L_z and the discrepancy of a factor of i will be explained soon. Now, we should be able to describe a finite rotation through an angle θ as an n -fold iteration of smaller rotations through an angle θ/n . As we take n larger, then θ/n becomes smaller and the approximation (4.47) for the smaller rotation through $\epsilon = \theta/n$ becomes better. Thus, we expect

$$R_z(\theta) = [R_z(\theta/n)]^n \approx \left(I + \frac{\theta L_z}{n} \right)^n$$

to become an equality in the limit $n \rightarrow \infty$. However, you should check, if it is not already a familiar fact, that

$$\lim_{n \rightarrow \infty} \left(I + \frac{X}{n} \right)^n = \sum_{n=0}^{\infty} \frac{X^n}{n!} = e^X \quad (4.48)$$

for any real number or matrix X . Thus, we can write

$$\boxed{R_z(\theta) = e^{\theta L_z}} \quad (4.49)$$

¹⁸See Hall [8] for further information and references.

where from here on out the exponential of a matrix or linear operator is *defined* by the power series in (4.48). Notice that the set $\{R_z(\theta) = e^{\theta L_z} \mid \theta \in \mathbb{R}\}$ is a subgroup of $SO(3)$; this can be seen by explicitly checking that

$$R_z(\theta_1)R_z(\theta_2) = R_z(\theta_1 + \theta_2), \quad (4.50)$$

or using the property¹⁹ of exponentials that $e^X e^Y = e^{X+Y}$, or recognizing intuitively that any two successive rotations about the z -axis yields another rotation about the z -axis. Notice that (4.50) says that if we consider $R_z(\theta)$ to be the image of θ under a map

$$R_z : \mathbb{R} \rightarrow SO(3), \quad (4.51)$$

then R_z is a homomorphism! Any such continuous homomorphism from the additive group \mathbb{R} to a matrix Lie group G is known as a **one-parameter subgroup**. One-parameter subgroups are actually very familiar to physicists; the set of rotations in \mathbb{R}^3 about *any* particular axis (not just the z -axis) is a one-parameter subgroup (where the parameter can be taken to be the rotation angle), as is the set of all boosts in a particular direction (in which case the parameter can be taken to be the absolute value u of the rapidity). We will also see that translations along a particular direction in both momentum and position space are one-parameter subgroups as well.

If we have a matrix X such that $e^{tX} \in G \ \forall t \in \mathbb{R}$, then the map

$$\begin{aligned} \exp : \mathbb{R} &\rightarrow G \\ t &\mapsto e^{tX} \end{aligned} \quad (4.52)$$

is a one-parameter subgroup, by the above-mentioned property of exponentials. Conversely, if we have a one-parameter subgroup $\gamma : \mathbb{R} \rightarrow G$, then we know that $\gamma(0) = I$ (since γ is a homomorphism), and, defining

$$X \equiv \frac{d\gamma}{dt}(0), \quad (4.53)$$

we have

$$\begin{aligned} \frac{d\gamma}{dt}(t) &= \lim_{h \rightarrow 0} \frac{\gamma(h+t) - \gamma(t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\gamma(h)\gamma(t) - \gamma(t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\gamma(h) - I}{h} \gamma(t) \\ &= \lim_{h \rightarrow 0} \frac{\gamma(h) - \gamma(0)}{h} \gamma(t) \\ &= X\gamma(t). \end{aligned}$$

¹⁹This is actually only true when X and Y commute; more on this later.

You should recall²⁰ that the first order linear matrix differential equation $\frac{d\gamma}{dt}(t) = X\gamma(t)$ has the unique solution $\gamma(t) = e^{tX}$. Thus, every one-parameter subgroup is of the form (4.52), so we have a one-to-one correspondence between one-parameter subgroups and matrices X such that $e^{tX} \in G \ \forall t \in \mathbb{R}$. The matrix X is sometimes said to ‘generate’ the corresponding one-parameter subgroup (e.g. L_z ‘generates’ rotations about the z -axis, according to (4.49)), and to each X there corresponds an ‘infinitesimal transformation’ $I + \epsilon X$. Really, though, X is best thought of as a derivative (by (4.53)), and (4.47) is best thought of as the first two terms of a Taylor expansion, as in

$$R_z(\theta) = R_z(0) + \theta \frac{dR_z}{d\theta}(0) + \dots \quad (4.54)$$

(you should verify that $\frac{dR_z}{d\theta}(0) = L_z$!). These matrices corresponding to one-parameter subgroups have many useful properties and their study leads to many interesting applications, both in mathematics and physics, so we give them a name: **Lie algebras**.

4.5 Lie Algebras—Definition, Properties, and Examples

Given a matrix Lie group $G \subset GL(n, \mathbb{C})$, we define its **Lie algebra** \mathfrak{g} to be the set of all matrices $X \in M_n(\mathbb{C})$ such that $e^{tX} \in G \ \forall t \in \mathbb{R}$. The first goal of this section is to establish some of the basic properties of Lie algebras. This will require a somewhat formal discussion, however, so we begin with a quick look at the Lie algebras of a couple of familiar Lie groups, so that you have examples to keep in mind later.

First, let us consider $GL(n, \mathbb{C})$ and its Lie algebra $\mathfrak{gl}(n, \mathbb{C})$. How can we describe $\mathfrak{gl}(n, \mathbb{C})$? If X is any element of $M_n(\mathbb{C})$, then e^{tX} is invertible for all t , and its inverse is merely e^{-tX} . Thus

$$\mathfrak{gl}(n, \mathbb{C}) = M_n(\mathbb{C}). \quad (4.55)$$

Likewise,

$$\mathfrak{gl}(n, \mathbb{R}) = M_n(\mathbb{R}). \quad (4.56)$$

Second, consider our good friend $O(3)$, whose Lie algebra is denoted $\mathfrak{o}(3)$. How can we describe $\mathfrak{o}(3)$? For any $X \in \mathfrak{o}(3)$, we have $e^{tX} \in O(3) \ \forall t \in \mathbb{R}$, which means that

$$(e^{tX})^T e^{tX} = e^{t(X^T)} e^{tX} = I \quad \forall t. \quad (4.57)$$

Differentiating this with respect to t yields

$$X^T e^{tX^T} e^{tX} + e^{tX^T} X e^{tX} = 0, \quad (4.58)$$

²⁰See any standard text on linear algebra and linear ODEs.

which we can evaluate at $t = 0$ to give

$$X^T + X = 0. \quad (4.59)$$

Thus, any $X \in \mathfrak{o}(3)$ must be antisymmetric! Will any real, antisymmetric X do? For any such X ,

$$\begin{aligned} (e^{tX})^T e^{tX} &= e^{t(X^T)} e^{tX} \\ &= e^{-tX} e^{tX} \\ &= I \end{aligned}$$

so $e^{tX} \in O(3) \forall t \in \mathbb{R}$, hence $X \in \mathfrak{o}(3)$. Thus, $\mathfrak{o}(3)$ is the set of *all* real antisymmetric matrices.

Now, $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{gl}(n, \mathbb{R})$, and $\mathfrak{o}(3)$ all share a few nice properties. First off, they are all real vector spaces, as you can easily check ($\mathfrak{gl}(n, \mathbb{C})$ can also be considered a complex vector space, but we will not consider it as such in this text). Secondly, they are *closed under commutators*, in the sense that if X and Y are elements of the Lie algebra, then so is

$$[X, Y] \equiv XY - YX. \quad (4.60)$$

This is trivial to verify in the case of $\mathfrak{gl}(n, \mathbb{C})$, and simple to verify in the case of $\mathfrak{o}(3)$. Thirdly, these Lie algebras (and, in fact, all sets of matrices) satisfy the **Jacobi Identity**,

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \forall X, Y, Z \in \mathfrak{g}. \quad (4.61)$$

This can be verified directly by expanding the commutators, and does not depend on any special properties of \mathfrak{g} , only on the definition of the commutator. We will have more to say about the Jacobi identity later.

The reason we have singled out these peculiar-seeming properties is that they will turn out to be important, and it turns out that *all* Lie algebras of matrix Lie groups enjoy them:

Proposition 4.1 *Let \mathfrak{g} be the Lie algebra of a matrix Lie group G . Then \mathfrak{g} is a real vector space, is closed under commutators, and all elements of \mathfrak{g} obey the Jacobi identity.*

Proving this turns out to be somewhat technical, so we will just sketch a proof here and refer you to Hall [8] for the details. Let \mathfrak{g} be the Lie algebra of a matrix Lie group G , and let $X, Y \in \mathfrak{g}$. We'd first like to show that $X + Y \in \mathfrak{g}$. Since \mathfrak{g} is closed under real scalar multiplication (why?), proving $X + Y \in \mathfrak{g}$ will be enough to show that \mathfrak{g} is a real vector space. The proof of this hinges on the following identity, known as the **Lie Product Formula**, which we state but do not prove:

$$e^{X+Y} = \lim_{m \rightarrow \infty} \left(e^{\frac{X}{m}} e^{\frac{Y}{m}} \right)^m. \quad (4.62)$$

This formula should be thought of as expressing the addition operation in \mathfrak{g} in terms of the product operation in G . With this in hand, we note that $A_m \equiv (e^{\frac{X}{m}} e^{\frac{Y}{m}})^m$ is a

convergent sequence, and that every term in the sequence is in G since it is a product of elements in G . Furthermore, the limit matrix $A = e^{X+Y}$ is in $GL(n, \mathbb{C})$ by (4.55). By the definition of a matrix Lie group, then, $A = e^{X+Y} \in G$, and thus \mathfrak{g} is a real vector space.

The second task is to show that \mathfrak{g} is closed under commutators. First, we claim that for any $X \in \mathfrak{gl}(n, \mathbb{C})$ and $A \in GL(n, \mathbb{C})$,

$$e^{AXA^{-1}} = Ae^XA^{-1}. \quad (4.63)$$

You can easily verify this by expanding the power series on both sides. This implies that if $X \in \mathfrak{g}$ and $A \in G$, then $AXA^{-1} \in \mathfrak{g}$ as well, since

$$e^{tAXA^{-1}} = Ae^{tX}A^{-1} \in G \quad \forall t \in \mathbb{R}. \quad (4.64)$$

Now let $A = e^{tY}$, $Y \in \mathfrak{g}$. Then $e^{tY}Xe^{-tY} \in \mathfrak{g} \quad \forall t$, and we can compute the derivative of this expression at $t = 0$:

$$\begin{aligned} \left. \frac{d}{dt} e^{tY}Xe^{-tY} \right|_{t=0} &= Y e^{tY}Xe^{-tY} \Big|_{t=0} - e^{tY}Xe^{-tY}Y \Big|_{t=0} \\ &= YX - XY. \end{aligned} \quad (4.65)$$

Since we also have

$$\left. \frac{d}{dt} e^{tY}Xe^{-tY} \right|_{t=0} = \lim_{h \rightarrow 0} \frac{e^{hY}Xe^{-hY} - X}{h} \quad (4.66)$$

and the right side is always in \mathfrak{g} since \mathfrak{g} is a vector space,²¹ this shows that $YX - XY = [Y, X]$ is in \mathfrak{g} .

The third and final task would be to verify the Jacobi identity, but as we pointed out above this holds for *any* set of matrices and can be easily verified by direct computation. Thus the proposition is established. \square

Before moving on to a discussion of specific Lie algebras, we should discuss the significance of the commutator. We proved above that all Lie algebras are closed under commutators, but so what? Why is this property worth singling out? Well, it turns out that the algebraic structure of the commutator on \mathfrak{g} is closely related to the algebraic structure of the product on G . This is most clearly manifested in the **Baker–Campbell–Hausdorff** (BCH) formula, which for X and Y sufficiently small²² expresses e^Xe^Y as a single exponential:

$$e^Xe^Y = e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+\cdots}. \quad (4.68)$$

²¹To be rigorous, we also need to note that \mathfrak{g} is closed in the topological sense, but this can be regarded as a technicality.

²²The size of a matrix $X \in M_n(\mathbb{C})$ is usually expressed by the **Hilbert–Schmidt norm**, defined as

$$\|X\| \equiv \sum_{i,j=1}^n |X_{ij}|^2. \quad (4.67)$$

If we introduce the basis $\{E_{ij}\}$ of $M_n(\mathbb{C})$ from Example 2.8, then we can identify $M_n(\mathbb{C})$ with \mathbb{C}^{n^2} , and then (4.67) is just the norm given by the standard Hermitian inner product.

It can be shown²³ that the series in the exponent converges for such X and Y , and that the series consists *entirely of iterated commutators*, so that the exponent really is an element of \mathfrak{g} (if the exponent had a term like XY in it, this would not be the case since matrix Lie algebras are in general not closed under ordinary matrix multiplication). Thus, the BCH formula plays a role analogous to that of the Lie product formula, but in the other direction: while the Lie product formula expresses Lie algebra addition in terms of group multiplication, **the BCH formula expresses group multiplication in terms of the commutator on the Lie algebra**. The BCH formula thus tells us that **much of the group structure of G is encoded in the structure of the commutator on \mathfrak{g}** . You will calculate the first few terms of the exponential in the BCH formula in Problem 4.11.

Now we are ready to look at the Lie algebras of some familiar matrix Lie groups.

Example 4.23 $\mathfrak{o}(n)$ and $\mathfrak{u}(n)$: *The Lie algebras of $O(n)$ and $U(n)$*

Recall from the beginning of this section that $\mathfrak{o}(3)$ is just the set of all real, anti-symmetric 3×3 matrices. You should note that the calculations which proved this did not depend in any way on n , so we can conclude in general that $\mathfrak{o}(n)$ is just the set of all real, antisymmetric $n \times n$ matrices. You can check that the matrices A_{ij} from Example 2.8 form a basis for $\mathfrak{o}(n)$, so that $\dim \mathfrak{o}(n) = \frac{n(n-1)}{2}$. As we will see in detail in the $n = 3$ case, the A_{ij} can be interpreted as generating rotations in the i - j plane.

As for $\mathfrak{u}(n)$, if we start with defining condition of $U(n)$, $A^\dagger A^{-1} = I$, and perform the same manipulations that led from (4.57) to (4.59), we find that for any $X \in \mathfrak{u}(n)$,

$$X^\dagger = -X. \quad (4.69)$$

In other words, X must be **anti-Hermitian**. Furthermore, the same argument that showed that any antisymmetric matrix is in $\mathfrak{o}(n)$ also shows that any anti-Hermitian matrix is in $\mathfrak{u}(n)$, so $\mathfrak{u}(n)$ is the set of *all* $n \times n$ anti-Hermitian matrices. In Exercise 2.5 you constructed a basis for $H_n(\mathbb{C})$, and as we mentioned in Example 2.4, multiplying a Hermitian matrix by i yields an anti-Hermitian matrix. Thus, if we multiply the basis from Exercise 2.5 by i we get a basis for $\mathfrak{u}(n)$, and it then follows that $\dim \mathfrak{u}(n) = n^2$. Note that a real antisymmetric matrix can also be considered an anti-Hermitian matrix, so that $\mathfrak{o}(n) \subset \mathfrak{u}(n)$. \square

Aside Before meeting some other Lie algebras, we have a loose end to tie up. We claimed earlier that most physical observables are best thought of as elements of Lie algebras (we will justify this statement in Sect. 4.7). However, we know that those observables must be represented by Hermitian operators (so that their eigenvalues are real), and yet here we see that the elements of $\mathfrak{o}(n)$ and $\mathfrak{u}(n)$ are *anti*-Hermitian! This is where our mysterious factor of i , which we mentioned in the last section, comes into play. Strictly speaking, $\mathfrak{o}(n)$ and $\mathfrak{u}(n)$ are real vector spaces whose elements are anti-Hermitian matrices. However, if we permit ourselves to ignore the real vector space structure and treat these matrices just as elements of $M_n(\mathbb{C})$, we can multiply them by i to get Hermitian matrices which we can then take to represent physical observables. In the physics literature, one

²³See Hall [8] for a nice discussion and Varadarajan [17] for a complete proof.

usually turns this around and *defines* generators of transformations to be these Hermitian observables, and then multiplies by i to get an anti-Hermitian operator which can then be exponentiated into an isometry. So one could say that the physicist's definition of the Lie algebra of a matrix Lie group G is $\mathfrak{g}_{\text{physics}} = \{X \in GL(n, \mathbb{C}) \mid e^{itX} \in G \ \forall t \in \mathbb{R}\}$. In the rest of this book we will stick with our original definition of the Lie algebra of a matrix Lie group, but you should be aware that the physicist's definition is often implicit in the physics literature.

Example 4.24 $\mathfrak{o}(n-1, 1)$: *The Lie algebra of $O(n-1, 1)$*

Let $X \in \mathfrak{o}(n-1, 1)$, the Lie algebra of the Lorentz group $O(n-1, 1)$. What do we know about X ? By the definition of a Lie algebra, and letting $[\eta] = \text{Diag}(1, 1, \dots, 1, -1)$, we have

$$e^{tX^T} [\eta] e^{tX} = [\eta] \quad \forall t.$$

Differentiating with respect to t and evaluating at $t = 0$ as before yields

$$X^T [\eta] + [\eta] X = 0. \quad (4.70)$$

Writing X out in block form with an $(n-1) \times (n-1)$ matrix X' for the spatial components (i.e. the X_{ij} where $i, j < n$) and vectors \mathbf{a} and \mathbf{b} for the components X_{in} and X_{ni} , $i < n$, the above equation reads

$$\begin{pmatrix} X'^T & -\mathbf{b} \\ \mathbf{a} & -X_{nn} \end{pmatrix} + \begin{pmatrix} X' & \mathbf{a} \\ -\mathbf{b} & -X_{nn} \end{pmatrix} = 0.$$

This implies that X has the form

$$X = \begin{pmatrix} X' & \mathbf{a} \\ \mathbf{a} & 0 \end{pmatrix}, \quad X' \in \mathfrak{o}(n-1), \quad \mathbf{a} \in \mathbb{R}^{n-1}.$$

One can think of X' as generating rotations in the $n-1$ spatial dimensions, and \mathbf{a} generating a boost along the direction it points in \mathbb{R}^{n-1} . We will discuss this in detail in the case $n = 4$ in the next section. \square

We have now described the Lie algebras of the isometry groups $O(n)$, $U(n)$, and $O(n-1, 1)$. What about their cousins $SO(n)$ and $SU(n)$? Since these groups are defined by the additional condition that they have unit determinant, examining their Lie algebras will require that we know how to evaluate the determinant of an exponential. This can be accomplished via the following beautiful and useful formula:

Proposition 4.2 *For any finite-dimensional matrix X ,*

$$\boxed{\det e^X = e^{\text{Tr } X}}. \quad (4.71)$$

A general proof of this is postponed to Problem 4.14, but we can gain some insight into the formula by proving it in the case where X is diagonalizable. Recall that *diagonalizable* means that there exists $A \in GL(n, \mathbb{C})$ such that

$$AXA^{-1} = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

where $\text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is a diagonal matrix with λ_i in the i th row and column, and zeros everywhere else. In this case, we have

$$\begin{aligned} \det e^X &= \det(Ae^XA^{-1}) \\ &= \det e^{AXA^{-1}} \end{aligned} \quad (4.72)$$

$$\begin{aligned} &= \det e^{\text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)} \\ &= \det \text{Diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}) \quad (\text{see exercise below}) \end{aligned} \quad (4.73)$$

$$= e^{\lambda_1} e^{\lambda_2} \dots e^{\lambda_n}$$

$$= e^{\lambda_1 + \lambda_2 + \dots + \lambda_n}$$

$$= e^{\text{Tr } X}$$

and so the formula holds. \square

Exercise 4.22 Prove (4.73). That is, use the definition of the matrix exponential as a power series to show that

$$e^{\text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n)} = \text{Diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}). \quad (4.74)$$

In addition to being a useful formula, Proposition 4.2 also provides a nice geometric interpretation of the Tr functional. Consider an arbitrary one-parameter subgroup $\{e^{tX}\} \subset GL(n, \mathbb{C})$. We have

$$\det e^{tX} = e^{\text{Tr } tX} = e^{t \text{Tr } X}, \quad (4.75)$$

so taking the derivative with respect to t and evaluating at $t = 0$ gives

$$\left. \frac{d}{dt} \det e^{tX} \right|_{t=0} = \text{Tr } X. \quad (4.76)$$

Since the determinant measures how an operator or matrix changes volumes (cf. Example 3.27), this tells us that the trace of the generator X gives the *rate at which volumes change* under the action of the corresponding one-parameter subgroup e^{tX} .

Example 4.25 $\mathfrak{so}(n)$ and $\mathfrak{su}(n)$, the Lie algebras of $SO(n)$ and $SU(n)$

Recall that $SO(n)$ and $SU(n)$ are the subgroups of $O(n)$ and $U(n)$ consisting of matrices with unit determinant. What additional condition must we impose on the generators X to ensure that $\det e^{tX} = 1 \ \forall t$? From (4.71), it is clear that $\det e^{tX} = 1 \ \forall t$ if and only if

$$\text{Tr } X = 0. \quad (4.77)$$

In the case of $\mathfrak{o}(n)$, this is actually already satisfied, since the antisymmetry condition implies that all the diagonal entries are zero. Thus, $\mathfrak{so}(n) = \mathfrak{o}(n)$, and both Lie algebras will henceforth be denoted as $\mathfrak{so}(n)$. We could have guessed that they would be equal; as discussed above, generators X are in one-to-one correspondence with ‘infinitesimal’ transformations $I + \epsilon X$, so the Lie algebra just tells us about

transformations that are close to the identity. However, the discussion from Example 4.12 generalizes easily to show that $O(n)$ is a group with two components, and that the component which contains the identity is just $SO(n)$. Thus the set of ‘infinitesimal’ transformations of both groups should be equal, and so should their Lie algebras. The same argument applies to $SO(3, 1)_o$ and $O(3, 1)$; their Lie algebras are thus identical, and will both be denoted as $\mathfrak{so}(3, 1)$.

For $\mathfrak{su}(n)$ the story is a little different. Here, the anti-Hermiticity condition only guarantees that the trace of an anti-Hermitian matrix is pure imaginary, so demanding that it actually be zero is a bona fide constraint. Thus, $\mathfrak{su}(n)$ can without redundancy be described as the set of traceless, anti-Hermitian $n \times n$ matrices. The tracelessness condition provides one additional constraint beyond anti-Hermiticity, so that

$$\dim \mathfrak{su}(n) = \dim \mathfrak{u}(n) - 1 = n^2 - 1. \quad (4.78)$$

Can you find a nice basis for $\mathfrak{su}(n)$?

Exercise 4.23 Show directly that $\mathfrak{so}(n)$ and $\mathfrak{su}(n)$ are vector spaces. Then prove the *cyclic property* of the Trace functional,

$$\mathrm{Tr}(A_1 A_2 \cdots A_n) = \mathrm{Tr}(A_2 \cdots A_n A_1), \quad A_i \in M_n(\mathbb{C}) \quad (4.79)$$

and use this to show directly that $\mathfrak{so}(n)$ and $\mathfrak{su}(n)$ are closed under commutators.

4.6 The Lie Algebras of Classical and Quantum Physics

We are now ready for a detailed look at the Lie algebras of the matrix Lie groups we discussed in Sect. 4.2.

Example 4.26 $\mathfrak{so}(2)$

As discussed in Example 4.23, $\mathfrak{so}(2)$ consists of all antisymmetric 2×2 matrices. All such matrices are of the form

$$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \quad (4.80)$$

and so $\mathfrak{so}(2)$ is one-dimensional and we may take as a basis

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (4.81)$$

You will explicitly compute that

$$e^{\theta X} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \theta \in \mathbb{R} \quad (4.82)$$

so that X really does generate counterclockwise rotations in the x - y plane.

Exercise 4.24 Verify (4.82) by explicitly summing the power series for $e^{\theta X}$. Also, consider the vector field X^\sharp induced by X on \mathbb{R}^2 , defined as follows: for every $\mathbf{r} \in \mathbb{R}^2$,

$$X^\sharp(\mathbf{r}) \equiv X\mathbf{r} \quad (4.83)$$

where X acts on \mathbf{r} by matrix multiplication. Sketch this vector field and show that the flow lines are circles. This gives direct geometric meaning to the matrix X , and is another sense in which ‘ X generates rotations’.

Example 4.27 $\mathfrak{so}(3)$

As remarked in Example 4.23, the matrices A_{ij} form a basis for $\mathfrak{o}(n)$. Specializing to $n = 3$ and renaming the A_{ij} as

$$L_x \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_z \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (4.84)$$

we can write an arbitrary element $X \in \mathfrak{so}(3)$ as

$$X = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} = xL_x + yL_y + zL_z \quad (4.85)$$

where you should recognize L_z from the previous section. Note that $\dim \mathfrak{so}(3) = 3$, which is also the number of parameters in $SO(3)$. You can check that the commutators of the basis elements work out to be

$$[L_x, L_y] = L_z$$

$$[L_y, L_z] = L_x$$

$$[L_z, L_x] = L_y$$

or, if we label the generators with numbers instead of letters,

$$[L_i, L_j] = \sum_{k=1}^3 \epsilon_{ijk} L_k.$$

(We will frequently change notation in this way when it proves convenient; hopefully this causes no confusion.) These are, of course, the well-known angular momentum commutation relations of quantum mechanics. The relation between the Lie algebra of $\mathfrak{so}(3)$ and the usual angular momentum operators (2.11) will be explained in the next chapter. Note that if we defined Hermitian $\mathfrak{so}(3)$ generators by $L'_i \equiv iL_i$, the commutation relations would become

$$[L'_i, L'_j] = \sum_{k=1}^3 i\epsilon_{ijk} L'_k,$$

which is the form found in most quantum mechanics texts.

We can think of the generator X in (4.85) as generating a rotation about the axis $[X]_{\{L_i\}} = (x, y, z)$, as follows: for any $v = (v_x, v_y, v_z) \in \mathbb{R}^3$, you can check that

$$Xv = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = (x, y, z) \times (v_x, v_y, v_z) = [X] \times v$$

so that if v lies along the axis $[X]$, then $Xv = 0$. This then implies that

$$e^{tX}[X] = [X]$$

(why?), so e^{tX} must be a rotation about $[X]$ (since it leaves $[X]$ fixed). In fact, if we take $[X]$ to be a unit vector and rename it $\hat{\mathbf{n}}$, and also rename t as θ , you will show below that

$$e^{\theta X} = \begin{pmatrix} n_x^2(1 - \cos \theta) + \cos \theta & n_x n_y(1 - \cos \theta) - n_z \sin \theta & n_x n_z(1 - \cos \theta) + n_y \sin \theta \\ n_y n_x(1 - \cos \theta) + n_z \sin \theta & n_y^2(1 - \cos \theta) + \cos \theta & n_y n_z(1 - \cos \theta) - n_x \sin \theta \\ n_z n_x(1 - \cos \theta) - n_y \sin \theta & n_z n_y(1 - \cos \theta) + n_x \sin \theta & n_z^2(1 - \cos \theta) + \cos \theta \end{pmatrix}, \quad (4.86)$$

proving our claim from Example 4.11 that this matrix represents a rotation about $\hat{\mathbf{n}}$ by an angle θ .

The astute reader may have noticed that we already made a connection between antisymmetric 3×3 matrices and rotations back in Example 3.30. There we saw that if $A(t)$ was a time-dependent orthogonal matrix representing the rotation of a rigid body, then the associated angular velocity bivector (in the space frame) was $[\tilde{\omega}] = \frac{dA}{dt} A^{-1}$. If we let $A(t)$ be a one-parameter subgroup $A(t) = e^{tX}$ generated by some $X \in \mathfrak{so}(3)$, then the associated angular velocity bivector is

$$\begin{aligned} [\tilde{\omega}] &= \frac{dA(t)}{dt} A^{-1} \\ &= X e^{tx} e^{-tx} \\ &= X. \end{aligned} \quad (4.87)$$

Thus **the angular velocity bivector is just the generator of the rotation!** Furthermore, applying the J map from Example 3.29 to both sides of (4.87) gives

$$[\omega] = [X]$$

and so the pseudovector ω is just the rotation generator expressed in coordinates $[X]$. Note that this agrees with our discussion above, where we found that $[X]$ gave the axis of rotation.

Exercise 4.25 Let $X \in \mathfrak{so}(3)$ be given by

$$X = \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix} \quad (4.88)$$

where $\hat{\mathbf{n}} = (n_x, n_y, n_z)$ is a unit vector. Verify (4.86) by explicitly summing the power series for $e^{\theta X}$.

Exercise 4.26 Using the basis $\mathcal{B} = \{L_x, L_y, L_z\}$ for $\mathfrak{so}(3)$, show that

$$[[X, Y]]_{\mathcal{B}} = [X]_{\mathcal{B}} \times [Y]_{\mathcal{B}}, \quad (4.89)$$

which shows that in components, the commutator is given by the usual cross product on \mathbb{R}^3 . (Note that on the left hand side of the above equation the two sets of brackets have entirely different meanings; one signifies the Lie bracket, while the other denotes the component representation of an element of the Lie algebra.)

Example 4.28 $\mathfrak{su}(2)$

We already met $\mathfrak{su}(2)$, the set of 2×2 traceless anti-Hermitian matrices, in Example 4.19. We took as a basis

$$S_x \equiv \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \quad S_y \equiv \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad S_z \equiv \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}. \quad (4.90)$$

Notice again that the number of parameters in the matrix Lie group $SU(2)$ is equal to the dimension of its Lie algebra $\mathfrak{su}(2)$. You can check that the commutation relation for these basis vectors is

$$[S_i, S_j] = \sum_{k=1}^3 \epsilon_{ijk} S_k \quad (4.91)$$

which is the same as the $\mathfrak{so}(3)$ commutation relations! Does that mean that $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are, in some sense, the same? And is this in some way related to the homomorphism from $SU(2)$ to $SO(3)$ that we discussed in Example 4.19? The answer to both these questions is yes, as we will discuss in the next few sections. We will also see that, just as $X = xL_x + yL_y + zL_z \in \mathfrak{so}(3)$ can be interpreted as generating a rotation about $(x, y, z) \in \mathbb{R}^3$, so can $Y = xS_1 + yS_2 + zS_3 \in \mathfrak{su}(2)$.

Exercise 4.27 Let $\hat{\mathbf{n}} = (n^1, n^2, n^3)$ be a unit vector. Show by direct calculation that

$$\begin{aligned} e^{\theta n^i S_i} &= \begin{pmatrix} \cos(\theta/2) - in_z \sin(\theta/2) & (-in_x - n_y) \sin(\theta/2) \\ (-in_x + n_y) \sin(\theta/2) & \cos(\theta/2) + in_z \sin(\theta/2) \end{pmatrix} \\ &= \cos(\theta/2)I + 2\sin(\theta/2)n^i S_i. \end{aligned} \quad (4.92)$$

You will use this formula in the problems to show that $e^{\theta n^i S_i}$ represents a rotation about the axis n by an angle θ .

Example 4.29 $\mathfrak{so}(3, 1)$

From Example 4.24, we know that an arbitrary $X \in \mathfrak{so}(3, 1)$ can be written as

$$X = \begin{pmatrix} X' & \mathbf{a} \\ \mathbf{a} & 0 \end{pmatrix}$$

with $X' \in \mathfrak{so}(3)$ and $\mathbf{a} \in \mathbb{R}^3$. Embedding the L_i of $\mathfrak{so}(3)$ into $\mathfrak{so}(3, 1)$ as

$$\tilde{L}_i \equiv \begin{pmatrix} L_i & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}$$

and defining new generators

$$K_1 \equiv \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad K_2 \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad K_3 \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

we have the following commutation relations (Exercise 4.28):

$$\begin{aligned} [\tilde{L}_i, \tilde{L}_j] &= \sum_{k=1}^3 \epsilon_{ijk} \tilde{L}_k \\ [\tilde{L}_i, K_j] &= \sum_{k=1}^3 \epsilon_{ijk} K_k \\ [K_i, K_j] &= - \sum_{k=1}^3 \epsilon_{ijk} \tilde{L}_k. \end{aligned} \tag{4.93}$$

As we mentioned in Example 4.24, the K_i can be interpreted as generating boosts along their corresponding axes. In fact, you will show in Exercise 4.28 that

$$e^{u^i K_i} = \begin{pmatrix} \frac{u_x^2(\cosh u - 1)}{u^2} + 1 & \frac{u_x u_y(\cosh u - 1)}{u^2} & \frac{u_x u_z(\cosh u - 1)}{u^2} & \frac{u_x}{u} \sinh u \\ \frac{u_y u_x(\cosh u - 1)}{u^2} & \frac{u_y^2(\cosh u - 1)}{u^2} + 1 & \frac{u_y u_z(\cosh u - 1)}{u^2} & \frac{u_y}{u} \sinh u \\ \frac{u_z u_x(\cosh u - 1)}{u^2} & \frac{u_z u_y(\cosh u - 1)}{u^2} & \frac{u_z^2(\cosh u - 1)}{u^2} + 1 & \frac{u_z}{u} \sinh u \\ \frac{u_x}{u} \sinh u & \frac{u_y}{u} \sinh u & \frac{u_z}{u} \sinh u & \cosh u \end{pmatrix}, \tag{4.94}$$

which as we know from Example 4.14 represents a boost in the direction of \mathbf{u} . Also, note that the last commutation relation in (4.93) says that if we perform two successive boosts in different directions, the *order* matters. Furthermore, the difference between performing them in one order and the other is actually a rotation!

Exercise 4.28 Verify the commutation relations (4.93), and verify (4.94) by explicitly summing the exponential power series.

Example 4.30 $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$

$\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is defined to be the Lie algebra of $SL(2, \mathbb{C})$, viewed as a *real* vector space. Since $SL(2, \mathbb{C})$ is just the set of all 2×2 complex matrices with unit determinant, $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is just the set of all *traceless* 2×2 complex matrices, and thus could be

viewed as a *complex* vector space, though we will not take that point of view here.²⁴ A (real) basis for $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is

$$\begin{aligned} S_1 &= \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ S_2 &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ S_3 &= \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\ \tilde{K}_1 &\equiv \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \tilde{K}_2 &\equiv \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \tilde{K}_3 &\equiv \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Note that $\tilde{K}_i = iS_i$. This fact simplifies certain computations, such as the ones you will perform in checking that these generators satisfy the following commutation relations:

$$\begin{aligned} [S_i, S_j] &= \sum_{k=1}^3 \epsilon_{ijk} S_k \\ [S_i, \tilde{K}_j] &= \sum_{k=1}^3 \epsilon_{ijk} \tilde{K}_k \\ [\tilde{K}_i, \tilde{K}_j] &= - \sum_{k=1}^3 \epsilon_{ijk} S_k. \end{aligned}$$

These are identical to the $\mathfrak{so}(3, 1)$ commutation relations! As in the case of $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$, this is intimately related to the homomorphism from $SL(2, \mathbb{C})$ to $SO(3, 1)_o$ that we described in Example 4.20. This will be discussed in Sect. 4.8.

Exercise 4.29 Check that the \tilde{K}_i generate boosts, as you would expect, by explicitly calculating $e^{u^i \tilde{K}_i}$ to get (4.35).

²⁴The space of all traceless 2×2 complex matrices viewed as a *complex* vector space is denoted $\mathfrak{sl}(2, \mathbb{C})$ without the \mathbb{R} subscript. In this case, the S_i suffice to form a basis. In this text, however, we will usually take Lie algebras to be real vector spaces, even if they naturally form complex vector spaces as well. You should be aware, though, that $\mathfrak{sl}(2, \mathbb{C})$ is fundamental in the theory of *complex* Lie algebras, and so in the math literature the Lie algebra of $SL(2, \mathbb{C})$ is almost always considered to be the complex vector space $\mathfrak{sl}(2, \mathbb{C})$ rather than the real Lie algebra $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$. We will have more to say about $\mathfrak{sl}(2, \mathbb{C})$ in the Appendix.

4.7 Abstract Lie Algebras

So far we have considered Lie algebras associated to matrix Lie groups, and we sketched proofs that these sets are real vector spaces which are closed under commutators. As in the case of abstract vector spaces and groups, however, we can now turn around and use these properties to *define* abstract Lie algebras. This will clarify the nature of the Lie algebras we have already met, as well as permit discussion of other examples relevant for physics.

That said, a (**real, abstract**) **Lie algebra** is defined to be a real vector space \mathfrak{g} equipped with a bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ called the **Lie bracket** which satisfies

1. $[X, Y] = -[Y, X] \quad \forall X, Y \in \mathfrak{g}$ (Antisymmetry)
2. $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \forall X, Y, Z \in \mathfrak{g}$ (Jacobi identity)

By construction, all Lie algebras of matrix Lie groups satisfy this definition (when we take the bracket to be the commutator), and we will see that it is precisely the above properties of the commutator that make those Lie algebras useful in applications. Furthermore, there are some (abstract) Lie algebras that arise in physics for which the bracket is *not* a commutator, and which are not usually associated with a matrix Lie group; this definition allows us to include those algebras in our discussion. We will meet a few of these algebras below, but first we consider two basic examples.

Example 4.31 $\mathfrak{gl}(V)$: *The Lie algebra of linear operators on a vector space*

Let V be a (possibly infinite-dimensional) vector space. We can turn $\mathcal{L}(V)$, the set of all linear operators on V , into a Lie algebra by taking the Lie bracket to be the commutator, i.e.

$$[T, U] \equiv TU - UT, \quad T, U \in \mathcal{L}(V). \quad (4.95)$$

Note that this is a commutator of *operators*, not matrices, though of course there is a nice correspondence between the two when V is finite-dimensional and we introduce a basis. This Lie bracket is obviously antisymmetric and can be seen to obey the Jacobi identity, so it turns $\mathcal{L}(V)$ into a Lie algebra which we will denote by $\mathfrak{gl}(V)$.²⁵ We will have more to say about $\mathfrak{gl}(V)$ as we progress.

Example 4.32 $\mathfrak{isom}(V)$: *The Lie algebra of anti-Hermitian operators*

Consider the setup of the previous example, except now let V be an inner product space. For any $T \in \mathcal{L}(V)$, the inner product on V allows us (via (4.9)) to define its adjoint T^\dagger , and we can then define $\mathfrak{isom}(V) \subset \mathfrak{gl}(V)$ to be the set of all **anti-Hermitian** operators, i.e. those which satisfy

²⁵There is a subtlety here: the vector space underlying $\mathfrak{gl}(V)$ is of course just $\mathcal{L}(V)$, so the difference between the two is just that one comes equipped with a Lie bracket, and the other is considered as a vector space with no additional structure.

$$\boxed{T^\dagger = -T.} \quad (4.96)$$

You can easily verify that $\text{isom}(V)$ is a Lie subalgebra of $\mathfrak{gl}(V)$. (A **Lie subalgebra** of a Lie algebra \mathfrak{g} is a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ that is also closed under the Lie bracket, and hence forms a Lie algebra itself.) If V is n -dimensional and we introduce an orthonormal basis, we see that the matrix form of $\text{isom}(V)$ is $\mathfrak{o}(n)$ if V is real and $\mathfrak{u}(n)$ if V is complex.

Example 4.33 *The Poisson bracket on phase space*

Consider a physical system with a $2n$ -dimensional phase space P parametrized by n generalized coordinates q_i and the n conjugate momenta p_i . The set of all complex-valued, infinitely differentiable²⁶ functions on P is a real vector space which we will denote by $\mathcal{C}(P)$. We can turn $\mathcal{C}(P)$ into a Lie algebra using the **Poisson bracket** as our Lie bracket, where the Poisson bracket is defined by

$$\{f, g\} \equiv \sum_i \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i}. \quad (4.97)$$

The antisymmetry of the Poisson bracket is clear, and the Jacobi identity can be verified directly by a brute-force calculation.

The functions in $\mathcal{C}(P)$ are known as **observables**, and the Poisson bracket thus turns the set of observables into one huge²⁷ Lie algebra. The standard quantization prescription, as developed by Dirac, Heisenberg, and the other founders of quantum mechanics, is to then interpret this Lie algebra of observables as a Lie subalgebra of $\mathfrak{gl}(\mathcal{H})$ for some Hilbert space \mathcal{H} (this identification is known as a Lie algebra **representation**, which is the subject of the next chapter). The commutator of the observables in $\mathfrak{gl}(\mathcal{H})$ is then just given by the Poisson bracket of the corresponding functions in $\mathcal{C}(P)$. Thus the set of all observables in quantum mechanics forms a Lie algebra, which is one of our main reasons for studying Lie algebras here.

Though $\mathcal{C}(P)$ is in general infinite-dimensional, it often has interesting finite-dimensional Lie subalgebras. For instance, if $P = \mathbb{R}^6$ and the q_i are just the usual cartesian coordinates for \mathbb{R}^3 , we can consider the three components of the angular momentum,

$$\begin{aligned} J_1 &= q_2 p_3 - q_3 p_2 \\ J_2 &= q_3 p_1 - q_1 p_3 \\ J_3 &= q_1 p_2 - q_2 p_1, \end{aligned} \quad (4.98)$$

²⁶A function is “infinitely differentiable” if it can be differentiated an arbitrary number of times. Besides the step function and its derivative, the Dirac delta ‘function’, most functions that one meets in classical physics and quantum mechanics are infinitely differentiable. This includes the exponential and trigonometric functions, as well as any other function that permits a power series expansion.

²⁷By this we mean infinite-dimensional, and usually requiring a basis that *cannot* be indexed by the integers but rather must be labeled by elements of \mathbb{R} or some other continuous set. You should recall from Sect. 3.7 that $L^2(\mathbb{R})$ was another such ‘huge’ vector space.

all of which are in $\mathcal{C}(\mathbb{R}^6)$. You will check below that the Poisson bracket of these functions turns out to be

$$\{J_i, J_j\} = \sum_{k=1}^3 \epsilon_{ijk} J_k, \quad (4.99)$$

which are of course the familiar angular momentum ($\mathfrak{so}(3)$) commutation relations (as mentioned above, this is in fact where the angular momentum commutation relations come from!). This then implies that $\mathfrak{so}(3)$ is a Lie subalgebra of $\mathcal{C}(\mathbb{R}^6)$. You may be wondering, however, why the angular momentum commutation relations are the same as the $\mathfrak{so}(3)$ commutation relations. What do the functions J_i have to do with generators of rotations? The answer has to do with a general relationship between symmetries and conserved quantities, which goes (roughly) as follows: Consider a classical system (i.e. a phase space P together with a Hamiltonian $H \in \mathcal{C}(P)$) which has a matrix Lie group G of canonical transformations²⁸ acting on it. If H is invariant²⁹ under the action of G , then G is said to be a group of **symmetries** of the system. In this case, one can then show³⁰ that for every $X \in \mathfrak{g}$ there is a function $f_X \in \mathcal{C}(P)$ which is *constant* along particle trajectories in P (where trajectories, of course, are given by solutions to Hamilton's equations). This fact is known as **Noether's theorem**, and it tells us that every element X of the Lie algebra gives a conserved quantity f_X . Furthermore, the Poisson bracket between two such functions f_X and f_Y is given just by the function associated to the Lie bracket of the corresponding elements of \mathfrak{g} , i.e.

$$\{f_X, f_Y\} = f_{[X, Y]}. \quad (4.100)$$

If $G = SO(3)$ acting on P by rotations, then it turns out³¹ that

$$f_X(q_i, p_i) = (\mathbf{J}(q_i, p_i), [X]), \quad X \in \mathfrak{so}(3) \quad (4.101)$$

where $\mathbf{J} = (J_1, J_2, J_3)$ is the usual angular momentum vector with components given by (4.98) and $(\cdot | \cdot)$ is the usual inner product on \mathbb{R}^3 . In particular, for $X = L_i \in \mathfrak{so}(3)$ we have

$$f_{L_i} = (\mathbf{J}, [L_i]) = (\mathbf{J}, e_i) = J_i. \quad (4.102)$$

Thus, the conserved quantity associated to rotations about the i th axis is just the i th component of angular momentum! This is the connection between angular momentum and rotations, and from (4.100) we see that the J_i *must* have the same commutation relations as the L_i , which is of course what we found in (4.99). \square

²⁸I.e. one-to-one and onto transformations which preserve the form of Hamilton's equations. See Goldstein [6].

²⁹Let $\phi_g : P \rightarrow P$ be the transformation of P corresponding to the group element $g \in G$. Then H is *invariant* under G if $H(\phi_g(p)) = H(p) \forall p \in P, g \in G$.

³⁰Under some mild assumptions. See Cannas [3] or Arnold [1], for example.

³¹See Arnold [1] for a discussion of Noether's theorem and a derivation of an equivalent formula.

Exercise 4.30 Verify (4.99). Also show that for a function $F \in \mathcal{C}(\mathbb{R}^6)$ that depends *only* on the coordinates q_i ,

$$\{p_j, F(q_i)\} = -\frac{\partial F}{\partial q_j} \quad (4.103)$$

and

$$\{J_3, F(q_i)\} = -q_1 \frac{\partial F}{\partial q_2} + q_2 \frac{\partial F}{\partial q_1} = -\frac{\partial F}{\partial \phi} \quad (4.104)$$

where ϕ is the azimuthal angle. We will interpret these results in the next chapter.

If we have a one-dimensional system with position coordinate q and conjugate momentum p , then $P = \mathbb{R}^2$ and $\mathcal{C}(P)$ contains another well-known Lie algebra: the Heisenberg algebra.

Example 4.34 *The Heisenberg algebra*

Define the **Heisenberg algebra** H to be the span of $\{q, p, 1\} \subset \mathcal{C}(\mathbb{R}^2)$, where $1 \in H$ is just the constant function with value 1. The only nontrivial Poisson bracket is between p and q , which you can check is just

$$\{q, p\} = 1, \quad (4.105)$$

which apart from the factors of \hbar (which we have dropped throughout the text) and i (which is just an artifact of the physicist's definition of a Lie algebra) is the familiar commutation relation from quantum mechanics. H is clearly closed under the Lie (Poisson) bracket, and is thus a Lie subalgebra of $\mathcal{C}(\mathbb{R}^2)$.

Can p and q be thought of as generators of specific transformations? Well, one of the most basic representations of p and q as operators is on the vector space $L^2(\mathbb{R})$, where

$$\hat{q}f(x) = xf(x) \quad (4.106)$$

$$\hat{p}f(x) = -\frac{df}{dx} \quad (4.107)$$

(again we drop the factor of i , in disagreement with the physicist's convention). Note that $[\hat{q}, \hat{p}] = \{q, p\} = 1$. If we exponentiate \hat{p} , we find (see Exercise 4.31 below)

$$e^{t\hat{p}}f(x) = f(x - t) \quad (4.108)$$

so that $\hat{p} = -\frac{d}{dx}$ generates translations along the x -axis! It follows that $\{e^{t\hat{p}} \mid t \in \mathbb{R}\}$ is the one-parameter subgroup of all translations along the x -axis. What about \hat{q} ? Well, if we work in the momentum representation of Example 3.17 so that we are dealing with the Fourier transform $\phi(p)$ of $f(x)$, you know from Exercise 3.20 that \hat{q} is represented by $i\frac{d}{dp}$, so if we multiply by i we get $i\hat{q} = -\frac{d}{dp}$ and thus

$$e^{it\hat{q}}\phi(p) = \phi(p - t), \quad (4.109)$$

which you can think of as a 'translation in momentum space'. We will explain the extra factor of i in the next chapter.

If we treat H as a complex vector space then we can consider another common basis for it, which is $\{Q, P, 1\} \subset \mathcal{C}(\mathbb{R}^2)$ where

$$Q \equiv \frac{p + iq}{\sqrt{2}}$$

$$P \equiv \frac{p - iq}{\sqrt{2}i}.$$

You can check that

$$\{Q, P\} = 1 \quad (4.110)$$

which you may recognize from classical mechanics as the condition that Q and P be *canonical variables* (see Goldstein [6]). Q and P are well suited to the solution of the one-dimensional harmonic oscillator problem, and you may recognize their formal similarity to the raising and lowering operators a and a^\dagger employed in the quantum-mechanical version of the same problem.

Q and P are not easily interpreted as generators of specific transformations, and our discussion of them helps explain why we defined *abstract* Lie algebras—so that we could work with spaces that behave *like* the Lie algebras of matrix Lie groups (in that they are vector spaces with a Lie bracket), but are not necessarily Lie algebras of matrix Lie groups themselves.

Exercise 4.31 Show by exponentiating $\hat{p} = -\frac{d}{dx}$ that $e^{t\hat{p}}f(x)$ is just the power series expansion for $f(x - t)$.

Exercise 4.32 Verify (4.110).

4.8 Homomorphism and Isomorphism Revisited

In Sect. 4.3 we used the notion of a group homomorphism to make precise the relationship between $SU(2)$ and $SO(3)$, as well as $SL(2, \mathbb{C})$ and $SO(3, 1)_o$. Now we will define the corresponding notion for Lie algebras to make precise the relationship between $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$, as well as $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ and $\mathfrak{so}(3, 1)$, and we will show how these relationships between Lie algebras arise as a consequence of the relationships between the corresponding groups.

That said, we define a **Lie algebra homomorphism** from a Lie algebra \mathfrak{g} to a Lie algebra \mathfrak{h} to be a linear map $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ that preserves the Lie bracket, in the sense that

$$\boxed{[\phi(X), \phi(Y)] = \phi([X, Y]) \quad \forall X, Y \in \mathfrak{g}.} \quad (4.111)$$

If ϕ is a vector space isomorphism (which implies that \mathfrak{g} and \mathfrak{h} have the same dimension), then ϕ is said to be a **Lie algebra isomorphism**. In this case, there is a one-to-one correspondence between \mathfrak{g} and \mathfrak{h} that preserves the bracket, so just as with group isomorphisms we consider \mathfrak{g} and \mathfrak{h} to be equivalent and write $\mathfrak{g} \simeq \mathfrak{h}$.

Sometimes the easiest way to prove that two Lie algebras \mathfrak{g} and \mathfrak{h} are isomorphic is with an astute choice of bases. Let $\{X_i\}_{i=1,\dots,n}$ and $\{Y_i\}_{i=1,\dots,n}$ be bases for \mathfrak{g} and \mathfrak{h} , respectively. Then the commutation relations take the form

$$[X_i, X_j] = \sum_{k=1}^n c_{ij}{}^k X_k$$

$$[Y_i, Y_j] = \sum_{k=1}^n d_{ij}{}^k Y_k$$

where the numbers $c_{ij}{}^k$ and $d_{ij}{}^k$ are known as the **structure constants** of \mathfrak{g} and \mathfrak{h} . (This is a bit of a misnomer, though, since the structure constants depend on a choice of basis, and are not as inherent a feature of the algebra as they make themselves out to be.) If one can exhibit bases such that $c_{ij}{}^k = d_{ij}{}^k \forall i, j, k$, then it is easy to check that the map

$$\phi : \mathfrak{g} \rightarrow \mathfrak{h}$$

$$v^i X_i \mapsto v^i Y_i$$

is a Lie algebra isomorphism. We will use this below to show that $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$ and $\mathfrak{so}(3, 1) \simeq \mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$.

Example 4.35 $\mathfrak{gl}(V)$ and $\mathfrak{gl}(n, C)$

Let V be an n -dimensional vector space over a set of scalars C . If we choose a basis for V , we can define a map from the Lie algebra $\mathfrak{gl}(V)$ of linear operators on V to the matrix Lie algebra $\mathfrak{gl}(n, C)$ by $T \mapsto [T]$. You can easily check that this is a Lie algebra isomorphism, so $\mathfrak{gl}(V) \simeq \mathfrak{gl}(n, C)$. If V is an inner product space we can restrict this isomorphism to $\mathfrak{isom}(V) \subset \mathfrak{gl}(V)$ to get

$$\mathfrak{isom}(V) \simeq \mathfrak{o}(n) \quad \text{if } C = \mathbb{R}$$

$$\mathfrak{isom}(V) \simeq \mathfrak{u}(n) \quad \text{if } C = \mathbb{C}.$$

Example 4.36 *The ad homomorphism*

Let \mathfrak{g} be a Lie algebra. We can use the bracket to turn $X \in \mathfrak{g}$ into a linear operator ad_X on \mathfrak{g} , by defining

$$\boxed{\text{ad}_X(Y) \equiv [X, Y], \quad X, Y \in \mathfrak{g}.} \quad (4.112)$$

We thus have a map

$$\text{ad} : \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$$

$$X \mapsto \text{ad}_X$$

between two Lie algebras. Is this map a Lie algebra homomorphism? It is if

$$\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y] \quad (4.113)$$

(notice that the bracket on the left is taken in \mathfrak{g} and on the right in $\mathfrak{gl}(\mathfrak{g})$). You will verify this in Exercise 4.33 below, where you will find that (4.113) is equivalent to the Jacobi identity! This is one way of interpreting the Jacobi identity: it guarantees that ad is a Lie algebra homomorphism for any Lie algebra \mathfrak{g} , and we will see that the ad homomorphism occurs frequently in physics. In fact, we already met it in Example 2.14! Besides its appearance in physics, the ad homomorphism is also fundamental in the beautiful structure theory of Lie algebras; this is treated in detail in Hall [8]. \square

Exercise 4.33 Verify that (4.113) is equivalent to the Jacobi identity.

Before we get to more physical examples, we need to explain how a continuous homomorphism from a matrix Lie group G to a matrix Lie group H leads to a Lie algebra homomorphism from \mathfrak{g} to \mathfrak{h} . This is accomplished by the following proposition:

Proposition 4.3 *Let $\Phi : G \rightarrow H$ be a continuous homomorphism from a matrix Lie group G to a matrix Lie group H . Then this induces a Lie algebra homomorphism $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ given by*

$$\phi(X) = \left. \frac{d}{dt} \Phi(e^{tX}) \right|_{t=0}. \quad (4.114)$$

Proof This proof is a little long, but is a nice application of all that we have been discussing. Let $\Phi : G \rightarrow H$ be such a homomorphism, and let $\{e^{tX}\}$ be a one-parameter subgroup in G . It's easy to see (check!) that $\{\Phi(e^{tX})\}$ is a one-parameter subgroup of H , and so

$$\left. \frac{d}{dt} \Phi(e^{tX}) \right|_{t=0} \equiv Z \quad (4.115)$$

is an element of \mathfrak{h} (by the definition of the Lie algebra of a matrix Lie group). We can thus define a map ϕ from \mathfrak{g} to \mathfrak{h} by $\phi(X) = Z$, and it follows that

$$\Phi(e^{tX}) = e^{t\phi(X)}. \quad (4.116)$$

We can roughly think of ϕ as the ‘infinitesimal’ version of Φ , with ϕ taking ‘infinitesimal’ transformations in \mathfrak{g} to ‘infinitesimal’ transformations in \mathfrak{h} . Is ϕ a Lie algebra homomorphism? There are several things to check. First, we must check that ϕ is linear, which we can prove by checking that $\phi(sX) = s\phi(X)$ and $\phi(X+Y) = \phi(X) + \phi(Y)$. To check that $\phi(sX) = s\phi(X)$, we can just differentiate using the chain rule:

$$\begin{aligned} \phi(sX) &= \left. \frac{d}{dt} \Phi(e^{tsX}) \right|_{t=0} \\ &= \left. \frac{d(st)}{dt} \frac{d}{d(st)} \Phi(e^{tsX}) \right|_{t=st=0} \end{aligned}$$

$$\begin{aligned}
&= sZ \\
&= s\phi(X).
\end{aligned}$$

Checking that $\phi(X + Y) = \phi(X) + \phi(Y)$ is a little more involved and involves a bit of calculation. The idea behind this calculation is to use the Lie product formula to express the addition in \mathfrak{g} in terms of the group product in G , then use the fact that Φ is a homomorphism to express this in terms of the product in H , and then use the Lie product formula in reverse to express this in terms of addition in \mathfrak{h} . We have

$$\begin{aligned}
\phi(X + Y) &= \frac{d}{dt} \Phi(e^{t(X+Y)}) \Big|_{t=0} \\
&= \frac{d}{dt} \Phi \left(\lim_{m \rightarrow \infty} (e^{\frac{tX}{m}} e^{\frac{tY}{m}})^m \right) \Big|_{t=0} && \text{by Lie Product formula} \\
&= \frac{d}{dt} \lim_{m \rightarrow \infty} \Phi \left((e^{\frac{tX}{m}} e^{\frac{tY}{m}})^m \right) \Big|_{t=0} && \text{since } \Phi \text{ is continuous} \\
&= \frac{d}{dt} \lim_{m \rightarrow \infty} (\Phi(e^{\frac{tX}{m}}) \Phi(e^{\frac{tY}{m}}))^m \Big|_{t=0} && \text{since } \Phi \text{ is a homomorphism} \\
&= \frac{d}{dt} \lim_{m \rightarrow \infty} (e^{\frac{t\phi(X)}{m}} e^{\frac{t\phi(Y)}{m}})^m \Big|_{t=0} && \text{by (4.116)} \\
&= \frac{d}{dt} e^{t(\phi(X) + \phi(Y))} \Big|_{t=0} && \text{by Lie product formula} \\
&= \phi(X) + \phi(Y).
\end{aligned}$$

So we have established that ϕ is a linear map from \mathfrak{g} to \mathfrak{h} . Does it preserve the bracket? Yes, but proving this also requires a bit of calculation. The idea behind this calculation is just to express everything in terms of one-parameter subgroups and then use the fact that Φ is a homomorphism. You can skip this calculation on a first reading, if desired. We have

$$\begin{aligned}
[\phi(X), \phi(Y)] &= \frac{d}{dt} e^{t\phi(X)} \phi(Y) e^{-t\phi(X)} && \text{by (4.65)} \\
&= \frac{d}{dt} \Phi(e^{tX}) \phi(Y) \Phi(e^{-tX}) && \text{by (4.116)} \\
&= \frac{d}{dt} \Phi(e^{tX}) \left(\frac{d}{ds} e^{s\phi(Y)} \right) \Phi(e^{-tX}) \\
&= \frac{d}{dt} \frac{d}{ds} \Phi(e^{tX}) \Phi(e^{sY}) \Phi(e^{-tX}) && \text{by (4.116)} \\
&= \frac{d}{dt} \frac{d}{ds} \Phi(e^{tX} e^{sY} e^{-tX}) && \text{since } \Phi \text{ is a homomorphism} \\
&= \frac{d}{dt} \frac{d}{ds} \Phi(e^{(se^{tX} Y e^{-tX})}) && \text{by (4.63)} \\
&= \frac{d}{dt} \phi(e^{tX} Y e^{-tX}) && \text{by definition of } \phi
\end{aligned}$$

$$\begin{aligned}
&= \phi\left(\frac{d}{dt}e^{tX}Ye^{-tX}\right) \quad \text{by Exercise 4.34 below} \\
&= \phi([X, Y]) \quad \text{by (4.65)}
\end{aligned}$$

and so ϕ is a Lie algebra homomorphism, and the proof is complete. \square

Exercise 4.34 Let ϕ be a linear map from a finite-dimensional vector space V to a finite-dimensional vector space W , and let $\gamma(t) : \mathbb{R} \rightarrow V$ be a differentiable V -valued function of t (you can think of this as a path in V parametrized by t). Show from the definition of a derivative that

$$\frac{d}{dt}\phi(\gamma(t)) = \phi\left(\frac{d}{dt}\gamma(t)\right) \quad \forall t. \quad (4.117)$$

Now let us make all this concrete by considering some examples.

Example 4.37 *$SU(2)$ and $SO(3)$ revisited*

Recall from Example 4.19 that we have a homomorphism $\rho : SU(2) \rightarrow SO(3)$ defined by the equation

$$[AXA^\dagger]_{\mathcal{B}} = \rho(A)[X]_{\mathcal{B}} \quad (4.118)$$

where $X \in \mathfrak{su}(2)$, $A \in SU(2)$ and $\mathcal{B} = \{S_x, S_y, S_z\}$. The induced Lie algebra homomorphism ϕ is given by

$$\phi(Y) = \left. \frac{d}{dt}\rho(e^{tY}) \right|_{t=0}, \quad (4.119)$$

and you will show below that this gives

$$\phi(S_i) = L_i, \quad i = 1-3. \quad (4.120)$$

This means that ϕ is one-to-one and onto, and since the commutation relations (i.e. structure constants) are the same for the S_i and L_i , we can then conclude from our earlier discussion that ϕ is a Lie algebra isomorphism, and thus $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$. You may have already known or guessed this, but the important thing to keep in mind is that this Lie algebra isomorphism actually stems from the *group* homomorphism between $SU(2)$ and $SO(3)$.

Exercise 4.35 Calculate $\left. \frac{d}{dt}\rho(e^{tS_i}) \right|_{t=0}$ and verify (4.120).

Example 4.38 *$SL(2, \mathbb{C})$ and $SO(3, 1)_o$ revisited*

Just as in the last example, we will now examine the Lie algebra isomorphism between $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ and $\mathfrak{so}(3, 1)$ that arises from the homomorphism $\rho : SL(2, \mathbb{C}) \rightarrow SO(3, 1)_o$. Recall that ρ was defined by

$$[AXA^\dagger]_{\mathcal{B}} = \rho(A)[X]_{\mathcal{B}} \quad (4.121)$$

where $A \in SL(2, \mathbb{C})$, $X \in H_2(\mathbb{C})$ and $\mathcal{B} = \{\sigma_x, \sigma_y, \sigma_z, I\}$. The induced Lie algebra homomorphism is given again by

$$\phi(Y) = \left. \frac{d}{dt} \rho(e^{tY}) \right|_{t=0}, \quad (4.122)$$

which, as you will again show, yields

$$\phi(S_i) = \tilde{L}_i \quad (4.123)$$

$$\phi(\tilde{K}_i) = K_i. \quad (4.124)$$

Thus ϕ is one-to-one, onto, and preserves the bracket (since the \tilde{L}_i and K_i have the same structure constants as the S_i and \tilde{K}_i), so $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}} \simeq \mathfrak{so}(3, 1)$. \square

Exercise 4.36 Verify (4.123) and (4.124).

What is the moral of the story from the previous two examples? How should one think about these groups and their relationships? Well, the homomorphism ρ allows us to interpret any $A \in SU(2)$ as a rotation and any $A \in SL(2, \mathbb{C})$ as a restricted Lorentz transformation. As we mentioned before, though, ρ is two-to-one, and in fact A and $-A$ in $SU(2)$ correspond to the same rotation in $SO(3)$, and likewise for $SL(2, \mathbb{C})$. However, A and $-A$ are not ‘close’ to each other; for instance, if we consider an infinitesimal transformation $A = I + \epsilon X$, we have $-A = -I - \epsilon X$, which is *not* close to the identity (though it is close to $-I$). Thus, the fact that ρ is not one-to-one cannot be discerned by examining the neighborhood around a given matrix; one has to look at the *global* structure of the group for that. So one might say that *locally*, $SU(2)$ and $SO(3)$ are identical, but *globally* they differ. In particular, they are identical when one looks at elements near the identity, which is why their Lie algebras are isomorphic. The same comments hold for $SL(2, \mathbb{C})$ and $SO(3, 1)_o$.

One important fact to take away from this is that **the correspondence between matrix Lie groups and Lie algebras is not one-to-one**; two different matrix Lie groups might have the same Lie algebra. Thus, if we start with a Lie algebra, there is no way to associate to it a unique matrix Lie group. This fact will have important implications in the next chapter.

Example 4.39 The Ad and ad homomorphisms

You may have found it curious that $\mathfrak{su}(2)$ was involved in the group homomorphism between $SU(2)$ and $SO(3)$. This is no accident, and Example 4.37 is actually an instance of a much more general construction which we now describe. Consider a matrix Lie group G and its Lie algebra \mathfrak{g} . We know that for any $A \in G$ and $X \in \mathfrak{g}$, AXA^{-1} is also in \mathfrak{g} , so we can actually define a linear operator Ad_A on \mathfrak{g} by

$$\boxed{\text{Ad}_A(X) = AXA^{-1}, \quad X \in \mathfrak{g}.} \quad (4.125)$$

We can think of Ad_A as the linear operator which takes a matrix X and applies the similarity transformation corresponding to A , as if A was implementing a change of basis. This actually allows us to define a homomorphism

$$\begin{aligned} \text{Ad} : G &\rightarrow GL(\mathfrak{g}) \\ A &\mapsto \text{Ad}_A, \end{aligned}$$

where you should quickly verify that $\text{Ad}_A \text{Ad}_B = \text{Ad}_{AB}$. Since Ad is a homomorphism between the two matrix Lie groups G and $GL(\mathfrak{g})$, we can consider the induced Lie algebra homomorphism $\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$. What does ϕ look like? Well, if $X \in \mathfrak{g}$ then $\phi(X) \in \mathfrak{gl}(\mathfrak{g})$ is the linear operator given by

$$\phi(X) = \left. \frac{d}{dt} \text{Ad}_{e^{tX}} \right|_{t=0}. \quad (4.126)$$

To figure out what this is, exactly, we evaluate the right hand side on $Y \in \mathfrak{g}$:

$$\begin{aligned} \phi(X)(Y) &= \left. \frac{d}{dt} (\text{Ad}_{e^{tX}}(Y)) \right|_{t=0} \\ &= \left. \frac{d}{dt} (e^{tX} Y e^{-tX}) \right|_{t=0} \\ &= [X, Y] \end{aligned}$$

so ϕ is nothing but the ad homomorphism of Example 4.36! Thus ad is the ‘infinitesimal’ version of Ad , and **the Lie bracket is the infinitesimal version of the similarity transformation**.

Note also that since Ad is a homomorphism, $\text{Ad}_{e^{tX}}$ is a one-parameter subgroup in $GL(\mathfrak{g})$, and its derivative at $t = 0$ is ad_X . From the discussion at the end of Sect. 4.4 we can then conclude that

$$\text{Ad}_{e^{tX}} = e^{t \text{ad}_X} \quad (4.127)$$

as an equality between operators on \mathfrak{g} . In other words, applying (4.127) to $Y \in \mathfrak{g}$,

$$e^{tX} Y e^{-tX} = Y + t[X, Y] + \frac{t^2}{2} [X, [X, Y]] + \frac{t^3}{3!} [X, [X, [X, Y]]] + \cdots \quad (4.128)$$

It is a nice exercise to expand the left hand side of this equation as a power series and verify the equality; this is Problem 4.13. \square

Exercise 4.37 Let $X, H \in M_n(\mathbb{C})$. Use (4.127) to show that

$$[X, H] = 0 \iff e^{tX} H e^{-tX} = H \quad \forall t \in \mathbb{R}. \quad (4.129)$$

If we think of H as a quantum-mechanical Hamiltonian, this shows how the invariance properties of the Hamiltonian (like invariance under rotations R) can be formulated in terms of commutators with the corresponding generators.

To make the connection between all this and Example 4.37, we note that usually Ad_A will preserve a metric on \mathfrak{g} (known as the **Killing Form** K ; see Problem 4.12), and thus

$$\text{Ad} : G \rightarrow \text{Isom}(\mathfrak{g}). \quad (4.130)$$

In the case of $G = SU(2)$ above, $\mathfrak{g} = \mathfrak{su}(2)$ is three-dimensional and K is positive-definite, so³²

³²Of course, this identification depends on a choice of basis, which was made when we chose to work with $\mathcal{B} = \{S_1, S_2, S_3\}$.

$$\text{Isom}(\mathfrak{su}(2)) \simeq O(3), \quad (4.131)$$

and thus $\text{Ad} : SU(2) \rightarrow O(3)$. You can check that this map is identical to the homomorphism ρ described in Example 4.19, and so we actually have $\text{Ad} : SU(2) \rightarrow SO(3)$! Thus the homomorphism between $SU(2)$ and $SO(3)$ is nothing but the Ad-joint map of $SU(2)$, where $\text{Ad}(g)$, $g \in SU(2)$ is orthogonal with respect to the Killing form on $\mathfrak{su}(2)$. We also have the corresponding Lie algebra homomorphism $\text{ad} : \mathfrak{su}(2) \rightarrow \mathfrak{so}(3)$, and we know that this must be equal to ϕ from Example 4.37, so $[\text{ad}_{S_i}]_{\mathcal{B}} = \phi(S_i) = L_i$.

Exercise 4.38 Using $[S_i, S_j] = \sum_{k=1}^3 \epsilon_{ijk} S_k$, compute the matrix representation of ad_{S_i} in the basis \mathcal{B} and verify explicitly that $[\text{ad}_{S_i}]_{\mathcal{B}} = L_i$.

4.9 Problems

4.1 In this problem we show that $SO(n)$ can be characterized as the set of all linear operators which take orthonormal (ON) bases into orthonormal bases and can be obtained continuously from the identity.

- (a) The easy part. Show that if a linear operator R takes ON bases into ON bases and is continuously obtainable from the identity, then $R \in SO(n)$. It's immediate that $R \in O(n)$; the trick here is showing that $\det R = 1$.
- (b) The converse. If $R \in SO(n)$, then it is immediate that R takes ON bases into ON bases. The slightly nontrivial part is showing that R is continuously obtainable from the identity. Prove this using induction, as follows: First, show that the claim is trivially true for $SO(1)$. Then suppose that the claim is true for $n - 1$. Take $R \in SO(n)$ and show that it can be continuously connected (via orthogonal similarity transformations) to a matrix of the form

$$\begin{pmatrix} 1 & \\ & R' \end{pmatrix}, \quad R' \in SO(n-1). \quad (4.132)$$

The claim then follows since by hypotheses R' can be continuously connected to the identity. (Hint: You will need the 2-D rotation which takes e_1 into Re_1 .)

4.2 In this problem we prove Euler's theorem that any $R \in SO(3)$ has an eigenvector with eigenvalue 1. This means that all vectors v proportional to this eigenvector are invariant under R , i.e. $Rv = v$, and so R fixes a line in space, known as the **axis of rotation**.

- (a) Show that λ being an eigenvalue of R is equivalent to $\det(R - \lambda I) = 0$. Refer to Problem 3.5 if necessary.
- (b) Prove Euler's theorem by showing that

$$\det(R - I) = 0. \quad (4.133)$$

Do this using the orthogonality condition and properties of the determinant. You should not have to work in components.

4.3 Show that the matrix (4.24) is just the component form (in the standard basis) of the linear operator

$$R(\hat{\mathbf{n}}, \theta) = L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}} + \cos \theta (I - L(\hat{\mathbf{n}}) \otimes \hat{\mathbf{n}}) + \sin \theta \hat{\mathbf{n}} \times \quad (4.134)$$

where (as you should recall) $L(v)(w) = (v, w)$ and the last term eats a vector v and spits out $\sin \theta \hat{\mathbf{n}} \times v$. Show that the first term is just the projection onto the axis of rotation $\hat{\mathbf{n}}$, and that the second and third term just give a counterclockwise rotation by θ in the plane perpendicular to $\hat{\mathbf{n}}$. Convince yourself that this is exactly what a rotation about $\hat{\mathbf{n}}$ should do.

4.4 Show that $SO(3, 1)_o$ is a subgroup of $O(3, 1)$. Remember that $SO(3, 1)_o$ is defined by 3 conditions: $|A| = 1$, $A_{44} > 1$, and (4.14). Proceed as follows:

- (a) Show that $I \in SO(3, 1)_o$.
- (b) Show that if $A \in SO(3, 1)_o$, then $A^{-1} \in SO(3, 1)_o$. Do this as follows:
 - (i) Verify that $|A^{-1}| = 1$.
 - (ii) Show that A^{-1} satisfies (4.14). Use this to deduce that A^T does also.
 - (iii) Write out the 44 component of (4.14) for both A and A^{-1} . You should get equations of the form

$$a_0^2 = 1 + \mathbf{a}^2 \quad (4.135)$$

$$b_0^2 = 1 + \mathbf{b}^2, \quad (4.136)$$

where $b_0 = (A^{-1})_{44}$. Clearly this implies $b_0 < -1$ or $b_0 > 1$. Now, write out the 44 component of the equation $AA^{-1} = I$. You should find

$$a_0 b_0 = 1 - \mathbf{a} \cdot \mathbf{b}. \quad (4.137)$$

If we let $a \equiv |\mathbf{a}|$, $b \equiv |\mathbf{b}|$ then the last equation implies

$$1 - ab < a_0 b_0 < 1 + ab. \quad (4.138)$$

Assume $b_0 < -1$ and use (4.136) to derive a contradiction to (4.138), hence showing that $b_0 = (A^{-1})_{44} > 1$, and that $A^{-1} \in SO(3, 1)_o$.

- (c) Show that if $A, B \in SO(3, 1)_o$, then $AB \in SO(3, 1)_o$. You may have to do some inequality manipulating to show that $(AB)_{44} > 0$.

4.5 In this problem we prove that any $A \in SO(3, 1)_o$ can be written as a product of a rotation and a boost.

- (a) If A is not a pure rotation then there is some relative velocity β between the standard reference frame and the new one described by A . Use the method of Exercise 4.11 to find β in terms of the components of A .
- (b) Let L be the pure boost of the form (4.29) corresponding to the β you found above. Show that $L^{-1}A$ is a rotation, by calculating that

$$\begin{aligned} (L^{-1}A)_{44} &= 1 \\ (L^{-1}A)_{i4} &= (L^{-1}A)_{4i} = 0, \quad i = 1, 2, 3. \end{aligned}$$

Conclude that $A = L(L^{-1}A)$ is the desired decomposition.

4.6 In this problem we show that any $A \in SL(2, \mathbb{C})$ can be decomposed as $A = \tilde{L}U$ where $U \in SU(2)$ and \tilde{L} is of the form (4.35). Unfortunately, it is a little too much work to prove this from scratch, so we will start with the **polar decomposition theorem**, which states that any $A \in SL(2, \mathbb{C})$ can be decomposed as $A = HU$ where $U \in SU(2)$ and H is Hermitian and **positive**, which means that $(v, Hv) > 0$ for all nonzero $v \in \mathbb{C}^2$ (here $(\cdot | \cdot)$ is the standard Hermitian inner product on \mathbb{C}^2). The polar decomposition theorem can be thought of as a higher-dimensional analog of the polar form of a complex number, $z = re^{i\theta}$. You will show below that the set of positive, Hermitian $H \in SL(2, \mathbb{C})$ is exactly the set of matrices of the form (4.35). This, combined with the polar decomposition theorem, yields the desired result. For more on the polar decomposition theorem itself, including a proof, see Hall [8].

(a) Show that an arbitrary Hermitian $1 \times H \in SL(2, \mathbb{C})$ can be written as

$$H = \begin{pmatrix} a+b & z \\ \bar{z} & a-b \end{pmatrix} \quad (4.139)$$

where

$$a, b \in \mathbb{R}, \quad z \in \mathbb{C}, \quad a^2 - b^2 - |z|^2 = 1. \quad (4.140)$$

(b) Show that any numbers a, b, z satisfying (4.140) can be written in the form

$$a = \pm \cosh u$$

$$b = v_z \sinh u$$

$$z = (v_x - i v_y) \sinh u$$

for some $u \in \mathbb{R}$ and unit vector $v \in \mathbb{R}^3$. Then show that positivity requires that $a = + \cosh u$. This puts H in the form (4.35).

(c) It remains to be shown that an H of the form (4.35) is actually positive. To do this, we employ a theorem (see Hoffman and Kunze [10]) which states that a matrix B is positive if and only if it is Hermitian and all its **principal minors** are positive, where its principal minors $\Delta_k(B)$, $k \leq n$ are the $k \times k$ partial determinants defined by

$$\Delta_k(B) \equiv \begin{vmatrix} B_{11} & \cdots & B_{1k} \\ \cdots & \cdots & \cdots \\ B_{k1} & \cdots & B_{kk} \end{vmatrix}. \quad (4.141)$$

Show that both the principal minors $\Delta_1(H)$ and $\Delta_2(H)$ are positive, so that H is positive.

4.7 In this problem we find an explicit formula for the map $\rho : SU(2) \rightarrow O(3)$ of Example 4.19 and use it to prove that ρ maps $SU(2)$ onto $SO(3)$ and has kernel $\pm I$.

- (a) Take an arbitrary $SU(2)$ matrix A of the form (4.25) and calculate AXA^\dagger for an arbitrary $X \in \mathfrak{su}(2)$ of the form (4.39).
- (b) Decomposing α and β into their real and imaginary parts, use (a) to compute the column vector $[AXA^\dagger]$.

- (c) Use (b) to compute the 3×3 matrix $\rho(A)$. Recall that $\rho(A)$ is defined by the equation $\rho(A)[X] = [AXA^\dagger]$.
- (d) Parametrize α and β as

$$\alpha = e^{i(\psi+\phi)/2} \cos \frac{\theta}{2}$$

$$\beta = ie^{i(\psi-\phi)/2} \sin \frac{\theta}{2}$$

as in (4.26). Substitute this into your expression for $\rho(A)$ and show that this is the transpose (or inverse, by the orthogonality relation) of (4.23).

- (e) Parametrize α and β as in (4.92) by

$$\alpha = \cos(\theta/2) - in_z \sin(\theta/2)$$

$$\beta = (-in_x - n_y) \sin(\theta/2)$$

and substitute this into your expression for $\rho(A)$ to get (4.24).

- (f) Conclude that $\det \rho(A) = 1$ and that ρ maps $SU(2)$ onto $SO(3)$. Also use your expression from (c) to show that the kernel of ρ is $\pm I$. (It's obvious that $\pm I \in K$. What takes a little calculation is showing that $\pm I$ is *all* of K .)

4.8 In this problem we find an explicit formula for the map $\rho : SL(2, \mathbb{C}) \rightarrow O(3, 1)$ of Example 4.20 and use it to prove that ρ maps $SL(2, \mathbb{C})$ onto $SO(3, 1)_o$ and has kernel $\pm I$. Note that since any $A \in SL(2, \mathbb{C})$ can be decomposed into $A = \tilde{L}\tilde{R}$ with $\tilde{R} \in SU(2)$ and \tilde{L} of the form (4.35), and any $B \in SO(3, 1)_o$ can be decomposed as $B = LR'$ where L is of the form (4.31) and $R' \in SO(3) \subset SO(3, 1)_o$, our task will be complete if we can show that $\rho(\tilde{L}) = L$. This we will do as follows:

- (a) To simplify computation, write the matrix (4.35) as

$$\tilde{L} = \begin{pmatrix} a+b & z \\ \bar{z} & a-b \end{pmatrix}, \quad a, b \in \mathbb{R}, \quad z \in \mathbb{C} \quad (4.142)$$

and calculate $\tilde{L}X\tilde{L}^\dagger$ for $X \in H_2(\mathbb{C})$ of the form (4.42).

- (b) Decomposing z into its real and imaginary parts, compute the column vector $[\tilde{L}X\tilde{L}^\dagger]$.
- (c) Use (b) to compute the 4×4 matrix $\rho(\tilde{L})$.
- (d) Substitute back in the original expressions for a , b and z

$$a = \cosh u/2$$

$$b = \frac{u_z}{u} \sinh u/2$$

$$z = \frac{1}{u}(u_x - iu_y) \sinh u/2$$

into $\rho(\tilde{L})$ to obtain (4.31).

4.9 In this problem we will prove the claim from Example 4.22 that, for a given permutation $\sigma \in S_n$, the number of transpositions in any decomposition of σ is either always odd or always even.

(a) Consider the polynomial

$$p(x_1, \dots, x_n) = \prod_{i < j} (x_i - x_j).$$

For example, for $n = 3$ and $n = 4$ this gives

$$\begin{aligned} p(x_1, x_2, x_3) &= (x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \\ p(x_1, x_2, x_3, x_4) &= (x_1 - x_2)(x_1 - x_3)(x_1 - x_4) \\ &\quad \times (x_2 - x_3)(x_2 - x_4)(x_3 - x_4). \end{aligned}$$

Define an action of $\sigma \in S_n$ on p by

$$(\sigma p)(x_1, \dots, x_n) \equiv p(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \prod_{i < j} (x_{\sigma(i)} - x_{\sigma(j)}).$$

Convince yourself that $\sigma p = \pm p$.

- (b) Let $\tau \in S_n$ be a transposition. Prove (or at least convince yourself) that $\tau p = -p$.
- (c) Now assume that σ has a decomposition into an even number of transpositions. Use p to prove that σ can then never have a decomposition into an *odd* number of transpositions. Use the same logic to show that if σ has a decomposition into an odd number of transpositions, then all of its decompositions must have an odd number of transpositions.

4.10 Consider the subset $H \subset GL(n, \mathbb{C})$ consisting of those matrices whose entries are real and rational. Show that H is in fact a subgroup, and construct a sequence of matrices in H that converge to an invertible matrix with *irrational* entries (there are many ways to do this!). This shows that H is a subgroup of $GL(n, \mathbb{C})$ which is *not* a matrix Lie group.

4.11 Let G be a Lie group and let $X, Y \in \mathfrak{g}$. Suppose that X and Y are small, so that $e^X e^Y$ is close to the identity and hence has a logarithm computable by the power series for \ln ,

$$\ln(X) = - \sum_{k=1}^{\infty} \frac{(I - X)^k}{k}. \quad (4.143)$$

By explicitly expanding out the relevant power series, show that up to third order in X and Y ,

$$\ln(e^X e^Y) = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \dots \quad (4.144)$$

Note that one would actually have to compute higher-order terms to verify that they can be written as commutators, and this gets very tedious. A more sophisticated proof is needed to show that every term in the series is an iterated commutator and hence an element of \mathfrak{g} . See Varadarajan [17] for such a proof.

4.12 Let G be a matrix Lie group. Its Lie algebra \mathfrak{g} comes equipped with a symmetric $(2, 0)$ tensor known as its **Killing Form**, denoted K and defined by

$$K(X, Y) \equiv -\text{Tr}(\text{ad}_X \text{ad}_Y). \quad (4.145)$$

(a) Show that K is Ad-invariant, in the sense that

$$K(\text{Ad}_A(X), \text{Ad}_A(Y)) = K(X, Y) \quad \forall X, Y \in \mathfrak{g}, A \in G.$$

(b) You know from Exercise 4.38 that $[\text{ad}_{S_i}] = L_i$. Use this to compute the components of K in the $\{S_i\}$ basis and prove that

$$[K] = 2I \quad (4.146)$$

so that K is positive-definite. This means that K is an inner product on $\mathfrak{su}(2)$, and so from part (a) we conclude that $\text{Ad}_A \in \text{Isom}(\mathfrak{su}(2)) \simeq O(3)$.

4.13 Prove directly that

$$\text{Ad}_{e^{tX}} = e^{t \text{ad}_X} \quad (4.147)$$

by induction, as follows: first verify that the terms first order in t on either side are equal. Then, assume that the n th order terms are equal (where n is an arbitrary integer), and use this to prove that the $n + 1$ th order terms are equal. Induction then shows that the terms of every order are equal, and so (4.147) is proven.

4.14 Prove the formula $e^{\text{Tr} X} = \det e^X$ as follows: Consider the determinant function as a homomorphism

$$\begin{aligned} \det : GL(n, \mathbb{C}) &\rightarrow \mathbb{C}^* \\ A &\mapsto \det A. \end{aligned}$$

This induces a Lie algebra homomorphism

$$\phi : \mathfrak{gl}(n, \mathbb{C}) \rightarrow \mathbb{C}. \quad (4.148)$$

We will show that $\phi(A) = \text{Tr} A$, which will imply $e^{\text{Tr} X} = \det e^X$ by (4.116). By definition, we have

$$\phi(X) = \left. \frac{d}{dt} \det(e^{tX}) \right|_{t=0}. \quad (4.149)$$

To evaluate this, expand the exponential to first order in t , plug into the determinant using the formula (3.84), and expand this to first order in t using properties of the determinant. This should yield the desired result.

Chapter 5

Basic Representation Theory

Now that we are familiar with groups and, in particular, the various transformation groups (i.e. matrix Lie groups) that arise in physics, we are ready to look at objects that ‘transform’ in specific ways under the action of these groups. The notion of an object ‘transforming’ in a specific way is made precise by the mathematical notion of a **representation**, which is essentially just a way of representing the elements of a group or Lie algebra as operators on a vector space (the objects which ‘transform’ are then just elements of the vector space). Representations are important in both classical and quantum physics: In classical physics, they clarify what we mean by a particular object’s ‘transformation properties’. In quantum mechanics, they both provide the basic mathematical framework (the Lie algebra of observables \mathfrak{g} acts on the Hilbert space at hand, so that the Hilbert space furnishes a representation of \mathfrak{g}) as well as clarify the notion of ‘vector’ and ‘tensor’ operators, which are usually introduced in a somewhat ad hoc way (much as we did in Sect. 3.7!). Furthermore, representation theory allows us to easily and generally prove many of the quantum mechanical ‘selection rules’ that are so handy in computation.

As with our discussion of tensors, we will treat mainly finite-dimensional vector spaces here but will occasionally be interested in infinite-dimensional applications. Rigorous treatment of these applications can be subtle and technical, though, so as before we will extend our results to certain infinite-dimensional cases without addressing the issues related to infinite-dimensionality. Again, you should be assured that all such applications are legitimate and can, in theory, be justified.

5.1 Representations: Definitions and Basic Examples

A **representation** of a group G is a vector space V together with a group homomorphism $\Pi : G \rightarrow GL(V)$. Sometimes they are written as a pair (Π, V) , though occasionally when the homomorphism Π is understood we will just talk about V , which is known as the **representation space**. If V is a real vector space then we say that (Π, V) is a **real** representation, and similarly if V is a complex vector space. We can think of the operators $\Pi(g)$ (which we will sometimes write as Π_g) as ‘representing’ the elements of G as invertible linear operators on V . If G is a matrix Lie group,

and V is finite-dimensional, and the group homomorphism $\Pi : G \rightarrow GL(V)$ is continuous, then Π induces a Lie algebra homomorphism $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ by (4.114). Any homomorphism from \mathfrak{g} to $\mathfrak{gl}(V)$ for some V is known as a **Lie algebra representation**, so every finite-dimensional representation of a Lie group G induces a representation of the corresponding Lie algebra \mathfrak{g} . The converse is not true, however; not every representation of \mathfrak{g} comes from a corresponding representation of G . This is intimately connected with the fact that for a given Lie algebra \mathfrak{g} , there is no unique matrix Lie group G that one can associate to it. We will discuss this in detail in the case of $\mathfrak{su}(2)$ and $\mathfrak{so}(3, 1)$ later.

In many of our physical applications the vector space V will come equipped with an inner product $(\cdot | \cdot)$ which is preserved by the operators Π_g (in other words, $\Pi(g) \in \text{Isom}(V) \forall g \in G$). In this case, we say that Π is a **unitary** representation, since each $\Pi(g)$ will be a unitary operator (cf. Example 4.5). If π is the induced Lie algebra representation, then $\forall X \in \mathfrak{g}$, we have

$$(v|w) = (\Pi(e^{tX})v | \Pi(e^{tX})w) \quad \forall v, w \in V, t \in \mathbb{R}. \quad (5.1)$$

Employing our standard trick of differentiating with respect to t and evaluating at $t = 0$ gives

$$\begin{aligned} 0 &= (\pi(X)v | w) + (v | \pi(X)w) \\ &= (\pi(X)v | w) + (\pi(X)^\dagger v | w) \\ &= ([\pi(X) + \pi(X)^\dagger]v | w) \\ &\Rightarrow \pi(X) + \pi(X)^\dagger = 0 \end{aligned}$$

so $\pi(X)$ is anti-Hermitian for all X ! Thus, if

$$\Pi : G \rightarrow \text{Isom}(V),$$

then

$$\pi : \mathfrak{g} \rightarrow \mathfrak{isom}(V) \quad (5.2)$$

as the notation suggests. Any Lie algebra homomorphism $\pi : \mathfrak{g} \rightarrow \mathfrak{isom}(V)$ is called a **unitary representation** of \mathfrak{g} , regardless of whether or not it comes from a unitary representation of G .

One of the reasons unitary representations are so common in physics is that in quantum mechanics, states are represented by unit vectors and so any linear operator T that takes a state into another state must preserve norms. This implies (as can be shown via an argument almost identical to the one used in Exercise 4.19) that T must be unitary, and so virtually all representations of interest in quantum mechanics are unitary.

Example 5.1 The trivial representation

As the name suggests, this example will be somewhat trivial, though we will end up referring to it later. For any group G (matrix or discrete) and vector space V , define the **trivial representation** of G on V by

$$\Pi(g) = I \quad \forall g \in G. \quad (5.3)$$

You will verify below that this is a representation. Suppose in addition that G is a matrix Lie group. What is the Lie algebra representation induced by Π ? For all $X \in \mathfrak{g}$ we have

$$\begin{aligned} \pi(X) &= \frac{d}{dt} \Pi(e^{tX}) \\ &= \frac{d}{dt} I \\ &= 0 \end{aligned}$$

so the trivial representation of a Lie algebra is given by $\pi(X) = 0 \quad \forall X \in \mathfrak{g}$.

Exercise 5.1 Let our representation space be $V = \mathbb{C}$. Show that $GL(\mathbb{C}) \simeq GL(1, \mathbb{C}) \simeq \mathbb{C}^*$ where \mathbb{C}^* is the group of nonzero complex numbers. Then verify that $\Pi : G \rightarrow \mathbb{C}^*$ as defined above is a group homomorphism, hence a representation. Also verify that $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathbb{C}) \simeq \mathbb{C}$ given by $\pi(X) = 0 \quad \forall X$ is a Lie algebra representation.

Example 5.2 The fundamental representation

Let G be a matrix Lie group. By definition, G is a subset of $GL(n, \mathbb{C}) = GL(C^n)$ for some n , so we can consider the obvious representation just given by interpreting the elements of G as operators (acting by matrix multiplication) on $V = C^n$. This is known as the **fundamental** (or **standard**) representation of G . If $G = O(3)$ or $SO(3)$ then $V = \mathbb{R}^3$ and the fundamental representation is known as the **vector** representation. If $G = SU(2)$ then $V = \mathbb{C}^2$ and the fundamental representation is known as the **spinor** representation. If $G = SO(3, 1)_o$ or $O(3, 1)$ then $V = \mathbb{R}^4$, and the fundamental representation is also known as the **vector** (or sometimes **four-vector**) representation. If $G = SL(2, \mathbb{C})$ then $V = \mathbb{C}^2$, and the fundamental representation is also known as the **spinor** representation. Vectors in this last representation are sometimes referred to more specifically as **left-handed spinors**, and are used to describe massless relativistic spin 1/2 particles like the neutrino.¹

Each of these group representations induces a representation of the corresponding Lie algebra which then goes by the same name, and which is also given just by interpreting the elements of $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ as linear operators. Since Lie algebras are vector spaces and a representation π is a linear map, we can describe any Lie algebra representation completely just by giving the image of the basis vectors under π (this is one of the nice features of Lie algebra representations; they are much easier to concretely visualize). Thus, the vector representation of $\mathfrak{so}(3)$ is given by

¹There is, of course, such a thing as a **right-handed spinor** as well, which we will meet in the next section and which is also used to describe massless spin 1/2 particles. The right and left-handed spinors are known collectively as **Weyl spinors**, in contrast to the **Dirac spinors**, which are used to describe *massive* spin 1/2 particles. We shall discuss Dirac spinors towards the end of this chapter.

$$\begin{aligned}\pi(L_x) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ \pi(L_y) &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \\ \pi(L_z) &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

where, of course, π is just the identity. Likewise, the spinor representation of $\mathfrak{su}(2)$ is given by

$$\begin{aligned}\pi(S_x) &= \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ \pi(S_y) &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \pi(S_z) &= \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\end{aligned}$$

and similarly for the vector representation of $\mathfrak{so}(3, 1)$ and the spinor representation of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$.

Exercise 5.2 Show that the fundamental representations of $SO(3)$, $O(3)$, and $SU(2)$ are unitary. (The fundamental representations of $SO(3, 1)_0$, $O(3, 1)$, and $SL(2, \mathbb{C})$ are *not* unitary, which can be guessed from the fact that the matrices in these groups are not unitary matrices. This stems from the fact that these groups preserve the Minkowski metric, which is not an inner product.)

Example 5.3 The adjoint representation

A less trivial class of examples is given by the Ad homomorphism of Example 4.39. Recall that Ad is a map from G to $GL(\mathfrak{g})$, where the operator Ad_A (for $A \in G$) is defined by

$$\text{Ad}_A(X) = AXA^{-1}, \quad X \in \mathfrak{g}.$$

In the context of representation theory, the Ad homomorphism is known as the **adjoint** representation (Ad, \mathfrak{g}). Note that the vector space of the adjoint representation is just the Lie algebra of G ! The adjoint representation is thus quite a natural construction, and is ubiquitous in representation theory (and elsewhere!) for that reason. To get a handle on what the adjoint representation looks like for some of the groups we have been working with, we consider the corresponding Lie algebra representation (ad, \mathfrak{g}), which you will recall acts as

$$\text{ad}_X(Y) = [X, Y], \quad X, Y \in \mathfrak{g}.$$

For $\mathfrak{so}(3)$ with basis $\mathcal{B} = \{L_i\}_{i=1-3}$, you have already calculated in Exercise 4.38 (using the isomorphic $\mathfrak{su}(2)$ with basis $\{S_i\}_{i=1-3}$) that

$$[\text{ad}_{L_i}]_{\mathcal{B}} = L_i$$

so for $\mathfrak{so}(3)$ the adjoint representation and fundamental representation are identical! (Note that we did not have to choose a basis when describing the fundamental representation because the use of the standard basis there is implicit.) Does this mean that the adjoint representations of the corresponding groups $SO(3)$ and $O(3)$ are also identical to the vector representation? Not quite. The adjoint representation of $SO(3)$ is identical to the vector representation (as we will show), but that does not carry over to $O(3)$; for $O(3)$, the inversion transformation $-I$ acts as minus the identity in the vector representation, but in the adjoint representation acts as

$$\text{Ad}_{-I}(X) = (-I)X(-I) = X \quad (5.4)$$

so Ad_{-I} is the identity! Thus the vector and adjoint representations of $O(3)$, though similar, are not identical, and so the adjoint representation is known as the **pseudovector** representation.² This will be discussed further in the next section.

What about the adjoint representations of $SU(2)$ and $\mathfrak{su}(2)$? Well, we already met these representations in Examples 4.19 and 4.37, and since $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$ and the adjoint representation of $\mathfrak{so}(3)$ is the vector representation, the adjoint representations of both $SU(2)$ and $\mathfrak{su}(2)$ are also known as their vector representations.

As for the adjoint representations of $SO(3, 1)_o$ and $O(3, 1)$, it is again useful to consider first the adjoint representation of their common Lie algebra, $\mathfrak{so}(3, 1)$. The vector space here is $\mathfrak{so}(3, 1)$ itself, which is six-dimensional and spanned by the basis $\mathcal{B} = \{\tilde{L}_i, K_j\}_{i,j=1-3}$. You will compute in Exercise 5.3 below that the matrix forms of $\text{ad}_{\tilde{L}_i}$ and ad_{K_i} are (in 3×3 block matrix form)

$$[\text{ad}_{\tilde{L}_i}] = \begin{pmatrix} L_i & 0 \\ 0 & L_i \end{pmatrix} \quad (5.5)$$

$$[\text{ad}_{K_i}] = \begin{pmatrix} 0 & -L_i \\ -L_i & 0 \end{pmatrix}. \quad (5.6)$$

From this we see that the \tilde{L}_i and K_i both transform like vectors under rotations ($\text{ad}_{\tilde{L}_i}$), but are mixed under boosts (ad_{K_i}). This is reminiscent of the behavior of the electric and magnetic field vectors, and it turns out (as we will see in Sect. 5.4) that the action of $\mathfrak{so}(3, 1)$ acting on itself via the adjoint representation is identical to the action of Lorentz transformation generators on the antisymmetric field tensor $F_{\mu\nu}$ from Example 3.16. The adjoint representation for $\mathfrak{so}(3, 1)$ is thus also known as the **antisymmetric second rank tensor** representation, as are the adjoint representations of the corresponding groups $SO(3, 1)_o$ and $O(3, 1)$. We omit a discussion of the adjoint representation of $SL(2, \mathbb{C})$ for technical reasons.³ \square

²You shouldn't be surprised at the nomenclature here, since we saw in Example 3.29 that pseudovectors are essentially 3×3 antisymmetric matrices, which are exactly $\mathfrak{so}(3)$!

³Namely, that the vector space in question, $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, is usually regarded as a three-dimensional complex vector space in the literature, not as a six-dimensional real vector space (which is the viewpoint of interest for us), so to avoid confusion we omit this topic. This will not affect any discussions of physical applications.

Exercise 5.3 Verify (5.5) and (5.6).

At this point your head may be swimming with nomenclature; the following table should help:

Group	Fundamental rep	Adjoint rep
$SO(3)$	vector	vector
$O(3)$	vector	pseudovector
$SU(2)$	spinor	vector
$SO(3, 1)_o$	vector	antisymmetric rank 2 tensor
$O(3, 1)$	vector	antisymmetric rank 2 tensor
$SL(2, \mathbb{C})$	Weyl spinor	—

5.2 Further Examples

The fundamental and adjoint representations of a matrix Lie group are the most basic examples of representations and are the ones out of which most others can be built, as we will see in the next section. There are, however, a few other representations of matrix Lie groups that you are probably already familiar with, as well as a few representations of abstract Lie algebras and discrete groups that are worth discussing. We will discuss these in this section, and return in the next section to developing the general theory.

Example 5.4 Representations of \mathbb{Z}_2

The notion of representations is useful not just for matrix Lie groups, but for more general groups as well. Consider the finite group $\mathbb{Z}_2 = \{1, -1\}$. For any vector space V , we can define the *alternating* representation (Π_{alt}, V) of \mathbb{Z}_2 by

$$\begin{aligned}\Pi_{\text{alt}}(1) &= I \\ \Pi_{\text{alt}}(-1) &= -I.\end{aligned}$$

This, along with the trivial representation, allows us to succinctly distinguish between the vector and pseudovector representations of $O(3)$ ⁴; when restricted to $\mathbb{Z}_2 \simeq \{I, -I\} \subset O(3)$, the fundamental (vector) representation of $O(3)$ becomes Π_{alt} , whereas the adjoint (pseudovector) representation becomes Π_{trivial} .

Another place where these representations of \mathbb{Z}_2 crop up is the theory of identical particles. Recall from Example 4.22 that there is a homomorphism $\text{sgn} : S_n \rightarrow \mathbb{Z}_2$ which tells us whether a given permutation is even or odd. If we then compose this map with either of the two representations introduced above, we get two representations of S_n : $\Pi_{\text{alt}} \circ \text{sgn}$, which is known as the *sgn* (read: ‘sign’) representation of S_n ,

⁴And, as we will see, between the vector and pseudovector representations of $O(3, 1)$.

and $\Pi_{\text{trivial}} \circ \text{sgn}$, which is just the trivial representation of S_n . If we consider an n -particle system with Hilbert space $\mathcal{T}_n^0(\mathcal{H})$, then from Example 4.22 we know that $S^n(\mathcal{H}) \subset \mathcal{T}_p^0(\mathcal{H})$ furnishes the trivial representation of S_n , whereas $\Lambda^n \mathcal{H} \subset \mathcal{T}_p^0(\mathcal{H})$ furnishes the sgn representation of S_n . This allows us to restate the symmetrization postulate once again, in its arguably most succinct form: *For a system composed of n identical particles, any state of the system lives either in the trivial representation of S_n (in which case the particles are known as bosons) or in the sgn representation of S_n (in which case the particles are known as fermions).*

Example 5.5 *The four-vector representation of $SL(2, \mathbb{C})$*

Recall from Example 4.20 that $SL(2, \mathbb{C})$ acts on $H_2(\mathbb{C})$ by sending $X \rightarrow AXA^\dagger$, where $X \in H_2(\mathbb{C})$ and $A \in SL(2, \mathbb{C})$. It is easy to see that this actually defines a representation $(\Pi, H_2(\mathbb{C}))$ given by

$$\Pi(A)(X) \equiv AXA^\dagger. \quad (5.7)$$

We already saw that if we take $\mathcal{B} = \{\sigma_x, \sigma_y, \sigma_z, I\}$ as a basis for $H_2(\mathbb{C})$, then $[\Pi(A)] \in SO(3, 1)$, so in this basis the action of $\Pi(A)$ looks like the action of restricted Lorentz transformations on four-vectors. Hence $(\Pi, H_2(\mathbb{C}))$ is also known as the **four-vector** representation of $SL(2, \mathbb{C})$.

Example 5.6 *The right-handed spinor representation of $SL(2, \mathbb{C})$*

As one might expect, the **right-handed spinor** representation $(\bar{\Pi}, \mathbb{C}^2)$ of $SL(2, \mathbb{C})$ is closely related to the fundamental (left-handed spinor) representation. It is defined simply by taking the adjoint inverse of the fundamental representation, that is,

$$\bar{\Pi}(A)v \equiv A^{\dagger-1}v, \quad A \in SL(2, \mathbb{C}), \quad v \in \mathbb{C}^2. \quad (5.8)$$

You can easily check that this defines a bona fide representation of $SL(2, \mathbb{C})$. The usual four component Dirac spinor can be thought of as a kind of ‘sum’ of a left-handed spinor and a right-handed one, as we will discuss later on. \square

The next few examples are instances of a general class of representations that are worth describing briefly. Say we are given a finite-dimensional vector space V , a representation Π of G on V , and a possibly infinite-dimensional vector space $C(V)$ of functions on V (this could be, for instance, the set $P_l(V)$ of all polynomial functions of a fixed degree l , or the set of all infinitely differentiable complex-valued functions $\mathcal{C}(V)$, or the set of all square-integrable functions $L^2(V)$). Then the representation Π on V induces a representation $\tilde{\Pi}$ on $C(V)$ as follows: if $f \in C(V)$, then the function $\tilde{\Pi}_g f \in C(V)$ is given by

$$(\tilde{\Pi}_g f)(v) \equiv f(\Pi_g^{-1}v), \quad g \in G, \quad f \in C(V), \quad v \in V. \quad (5.9)$$

There are a couple things to check. First, it must be verified that $\tilde{\Pi}_g f$ is actually an element of $C(V)$; this, of course, depends on the exact nature of $C(V)$ and of Π and must be checked independently for each example. Assuming this is true, we also

need to verify that $\tilde{\Pi}(g)$ is a linear operator and that it satisfies $\tilde{\Pi}(gh) = \tilde{\Pi}(g)\tilde{\Pi}(h)$. This computation is the same for all such examples, and you will perform it in Exercise 5.4 below.

Exercise 5.4 Confirm that if (Π, V) is a representation of G , then $\tilde{\Pi}_g$ is a linear operator on $C(V)$, and that $\tilde{\Pi}(g_1)\tilde{\Pi}(g_2) = \tilde{\Pi}(g_1g_2)$. What happens if you try to define $\tilde{\Pi}$ using g instead of g^{-1} ?

Example 5.7 *The spin s representation of $\mathfrak{su}(2)$ and polynomials on \mathbb{C}^2*

We already know that the Hilbert space corresponding to a spin s particle fixed in space is \mathbb{C}^{2s+1} . Since the spin angular momentum \mathbf{S} is an observable for this system and the S_i have the $\mathfrak{su}(2)$ structure constants, \mathbb{C}^{2s+1} must be a representation of $\mathfrak{su}(2)$. How is this representation defined? Usually one answers this question by considering the eigenvalues of S_z and showing that for any finite-dimensional representation V , the eigenvalues of S_z lie between $-s$ and s for some half-integral s . V is then defined to be the span of the eigenvectors of S_z (from which we conclude that $\dim V = 2s + 1$), and the action of S_x and S_y is determined by the $\mathfrak{su}(2)$ commutation relations. This construction is important and will be presented in Sect. 5.8, but it is also rather abstract; for the time being we present an alternate construction which is more concrete.

Consider the set of all degree l polynomials on \mathbb{C}^2 , i.e. the set of all degree l polynomials in the two complex variables z_1 and z_2 . This is a complex vector space, denoted $P_l(\mathbb{C}^2)$, and has basis

$$\mathcal{B}_l = \{z_1^{l-k}z_2^k \mid 0 \leq k \leq l\} = \{z_1^l, z_1^{l-1}z_2, \dots, z_1z_2^{l-1}, z_2^l\}$$

and hence dimension $l + 1$. The fundamental representation of $SU(2)$ on \mathbb{C}^2 then induces an $SU(2)$ representation $(\Pi_l, P_l(\mathbb{C}^2))$ as described above: given a polynomial function $p \in P_l(\mathbb{C}^2)$ and $A \in SU(2)$, $\Pi_l(A)(p)$ is the degree l polynomial given by

$$(\Pi_l(A)(p))(v) \equiv p(A^{-1}v), \quad v \in \mathbb{C}^2.$$

To make this representation concrete, consider the degree one polynomial $p(z_1, z_2) = z_1$. This polynomial function just picks off the first coordinate of any $v \in \mathbb{C}^2$. Let

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$$

so that

$$A^{-1} = \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix}.$$

Then

$$(\Pi_1(A)p)(v) = p(A^{-1}v)$$

but

$$A^{-1}v = \begin{pmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \bar{\alpha}z_1 - \beta z_2 \\ \bar{\beta}z_1 + \alpha z_2 \end{pmatrix} \quad (5.10)$$

so then

$$\Pi_1(A)z_1 = \bar{\alpha}z_1 - \beta z_2.$$

Likewise, (5.10) tells us that

$$\Pi_1(A)z_2 = \bar{\beta}z_1 + \alpha z_2.$$

From this, the action of $\Pi_l(A)$ on higher order polynomials can be determined since $\Pi_l(A)(z_1^{l-k}z_2^k) = (\Pi_1(A)z_1)^{l-k}(\Pi_1(A)z_2)^k$.

We can then consider the induced Lie algebra representation $(\pi_l, P_l(\mathbb{C}^2))$, in which (as you will show) $z_1^{l-k}z_2^k$ is an eigenvector of S_z with eigenvalue $i(\frac{l}{2} - k)$, so that

$$[S_z]_{\mathcal{B}_l} = i \begin{pmatrix} l/2 & & & \\ & l/2 - 1 & & \\ & & \dots & \\ & & & -l/2 \end{pmatrix}. \quad (5.11)$$

If we let $s \equiv l/2$ then we recognize this as the usual form (up to that pesky factor of i) of S_z acting on the Hilbert space of a spin s particle. \square

Exercise 5.5 Use the definition of induced Lie algebra representations and the explicit form of e^{tS_i} in the fundamental representation (which can be deduced from (4.92)) to compute the action of the operators $\pi_1(S_i)$ on the functions z_1 and z_2 in $P_1(\mathbb{C}^2)$. Show that we can write these operators in differential form as

$$\begin{aligned} \pi_1(S_1) &= \frac{i}{2} \left(z_2 \frac{\partial}{\partial z_1} + z_1 \frac{\partial}{\partial z_2} \right) \\ \pi_1(S_2) &= \frac{1}{2} \left(z_2 \frac{\partial}{\partial z_1} - z_1 \frac{\partial}{\partial z_2} \right) \\ \pi_1(S_3) &= \frac{i}{2} \left(z_1 \frac{\partial}{\partial z_1} - z_2 \frac{\partial}{\partial z_2} \right). \end{aligned} \quad (5.12)$$

Prove that these expressions also hold for the operators $\pi_l(S_i)$ on $P_l(\mathbb{C}^2)$. Verify the $\mathfrak{su}(2)$ commutation relations directly from these expressions. Finally, use (5.12) to show that

$$(\pi_l(S_z))(z_1^{l-k}z_2^k) = i(l/2 - k)z_1^{l-k}z_2^k.$$

Example 5.8 $L^2(\mathbb{R}^3)$ as a representation of $SO(3)$

Recall from Example 3.18 that $L^2(\mathbb{R}^3)$ is the set of all complex-valued square-integrable functions on \mathbb{R}^3 , and is physically interpreted as the Hilbert space of a spinless particle moving in three dimensions. As in the previous example, the fundamental representation of $SO(3)$ on \mathbb{R}^3 induces an $SO(3)$ representation $(\Pi, L^2(\mathbb{R}^3))$ by

$$(\Pi_R(f))(\mathbf{x}) \equiv f(R^{-1}\mathbf{x}), \quad f \in L^2(\mathbb{R}^3), \quad R \in SO(3), \quad \mathbf{x} \in \mathbb{R}^3. \quad (5.13)$$

One can think of $\Pi(R)f$ as just a “rotated” version of the function f . This representation is unitary, as we will now digress for a moment to show.

Let $(\cdot | \cdot)$ denote the Hilbert space inner product on $L^2(\mathbb{R}^3)$. To show that Π is unitary, we need to show that

$$(\Pi(R)f | \Pi(R)g) = (f | g) \quad \forall f, g \in L^2(\mathbb{R}^3). \quad (5.14)$$

First off, we have

$$\begin{aligned} (\Pi(R)f | \Pi(R)g) &= \int d^3x (\Pi_R \tilde{f})(\mathbf{x}) (\Pi_R g)(\mathbf{x}) \\ &= \int d^3x \tilde{f}(R^{-1}\mathbf{x}) g(R^{-1}\mathbf{x}). \end{aligned}$$

If we now change variables to $\mathbf{x}' = R^{-1}\mathbf{x}$ and remember to include the Jacobian, we get

$$\begin{aligned} (\Pi(R)f | \Pi(R)g) &= \int d^3x' \left| \frac{\partial(x, y, z)}{\partial(x', y', z')} \right| \tilde{f}(\mathbf{x}') g(\mathbf{x}') \\ &= \int d^3x' \tilde{f}(\mathbf{x}') g(\mathbf{x}') \\ &= (f | g) \end{aligned}$$

where you will verify below that the Jacobian determinant $|\frac{\partial(x, y, z)}{\partial(x', y', z')}|$ is equal to one. This should be no surprise; the Jacobian tells us how volumes change under a change of variables, but since in this case the change of variables is given just by a rotation (which we know preserves volumes), we should expect the Jacobian to be one.

Though the definition of Π might look strange, it is actually the action of rotations on position kets that we are familiar with; if we act on the basis ‘vector’ $|\mathbf{x}_0\rangle = \delta(\mathbf{x} - \mathbf{x}_0)$, we find that (for arbitrary $\psi \in L^2(\mathbb{R}^3)$)

$$\begin{aligned} \langle \psi | \Pi_R | \mathbf{x}_0 \rangle &= \int d^3x \tilde{\psi}(\mathbf{x}) \delta(R^{-1}\mathbf{x} - \mathbf{x}_0) \\ &= \int d^3x' \tilde{\psi}(R\mathbf{x}') \delta(\mathbf{x}' - \mathbf{x}_0) \quad \text{where we let } \mathbf{x}' \equiv R^{-1}\mathbf{x} \\ &= \tilde{\psi}(R\mathbf{x}_0) \end{aligned}$$

hence we must have

$$\Pi_R |\mathbf{x}_0\rangle = |R\mathbf{x}_0\rangle$$

which is the familiar action of rotations on position kets.

What does the corresponding representation of $\mathfrak{so}(3)$ look like?⁵ As mentioned above, we can get a handle on that just by computing $\pi(L_i)$, since π is a linear map

⁵We should mention here that the infinite-dimensionality of $L^2(\mathbb{R}^3)$ makes a proper treatment of the induced Lie algebra representation quite subtle; for instance, we calculate in this example that the elements of $\mathfrak{so}(3)$ are to be represented by differential operators, yet not all functions in $L^2(\mathbb{R}^3)$ are differentiable! (Just think of a step function which is equal to 1 inside the unit sphere and 0 outside the unit sphere; this function is not differentiable at $r = 1$.) In this example and elsewhere, we ignore such subtleties.

and any $X \in \mathfrak{so}(3)$ can just be expressed as a linear combination of the L_i . Hence we compute

$$\begin{aligned}
 (\pi_{L_i} f)(\mathbf{x}) &= \left. \frac{d}{dt} (\Pi_{e^{tL_i}} f)(\mathbf{x}) \right|_{t=0} && \text{by the def. of } \pi \\
 &= \left. \frac{d}{dt} f(e^{-tL_i} \mathbf{x}) \right|_{t=0} && \text{by the def. of } \Pi \\
 &= \sum_{j=1}^3 \frac{\partial f}{\partial x^j}(\mathbf{x}) \left. \frac{d}{dt} (e^{-tL_i} \mathbf{x})^j \right|_{t=0} && \text{by the multivariable chain rule} \\
 &= \sum_{j=1}^3 \frac{\partial f}{\partial x^j}(\mathbf{x}) (-L_i \mathbf{x})^j \\
 &= - \sum_{j,k=1}^3 \frac{\partial f}{\partial x^j}(\mathbf{x}) (L_i)_{jk} x^k \\
 &= \sum_{j,k=1}^3 \epsilon_{ijk} x^k \frac{\partial f}{\partial x^j}(\mathbf{x}) && \text{since } (L_i)_{jk} = -\epsilon_{ijk}
 \end{aligned}$$

so that, relabeling dummy indices,

$$\pi(L_i) = - \sum_{j,k=1}^3 \epsilon_{ijk} x^j \frac{\partial}{\partial x^k}.$$

More concretely, we have

$$\boxed{
 \begin{aligned}
 \pi(L_x) &= z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z} \\
 \pi(L_y) &= x \frac{\partial}{\partial z} - z \frac{\partial}{\partial x} \\
 \pi(L_z) &= y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}
 \end{aligned}
 } \tag{5.15}$$

which, up to our usual factor of i , is just (2.11)! □

Exercise 5.6 Verify that if $\mathbf{x}' = R^{-1} \mathbf{x}$ for some rotation $R \in SO(3)$, then

$$\left| \frac{\partial(x, y, z)}{\partial(x', y', z')} \right| = 1. \tag{5.16}$$

Example 5.9 $\mathcal{H}_l(\mathbb{R}^3)$, $\tilde{\mathcal{H}}_l$ and $L^2(S^2)$ as representations of $SO(3)$

Recall from Chap. 2 that $\mathcal{H}_l(\mathbb{R}^3)$ is the vector space of all harmonic complex-valued degree l polynomials on \mathbb{R}^3 . Since $\mathcal{H}_l(\mathbb{R}^3)$ is a space of functions on \mathbb{R}^3 , we get an $SO(3)$ representation of the same form as in the last example, namely

$$(\Pi_l(R)(f))(\mathbf{x}) \equiv f(R^{-1}\mathbf{x}), \quad f \in \mathcal{H}_l(\mathbb{R}^3), \quad R \in SO(3), \quad \mathbf{x} \in \mathbb{R}^3.$$

You will check below that if f is a harmonic polynomial of degree l , then $\Pi_l(R)f$ is too, so that Π_l really is a representation on $\mathcal{H}_l(\mathbb{R}^3)$. The induced $\mathfrak{so}(3)$ representation also has the same form as in the previous example.

If we now restrict all the functions in $\mathcal{H}_l(\mathbb{R}^3)$ to the unit sphere, we get a representation of $SO(3)$ on $\tilde{\mathcal{H}}_l$, the space of spherical harmonics of degree l . Concretely, we can describe this representation by writing $Y(\theta, \phi)$ as $Y(\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is a unit vector giving the point on the sphere which corresponds to (θ, ϕ) . Then the $SO(3)$ representation $(\tilde{\Pi}, \tilde{\mathcal{H}}_l)$ is given simply by

$$(\tilde{\Pi}_l(R)Y)(\hat{\mathbf{n}}) \equiv Y(R^{-1}\hat{\mathbf{n}}). \quad (5.17)$$

The interesting thing about this representation is that it also turns out to be unitary! The inner product in this case is just given by integration over the sphere with the usual area form, i.e.

$$(Y_1(\theta, \phi)|Y_2(\theta, \phi)) \equiv \int_0^\pi \int_0^{2\pi} \bar{Y}_1(\theta, \phi) Y_2(\theta, \phi) \sin \theta \, d\phi \, d\theta. \quad (5.18)$$

Proving that (5.17) is unitary with respect to this inner product is straightforward but tedious, so we omit the calculation.⁶

One nice thing about this inner product, though, is that we can use it to define a notion of square-integrability just as we did for \mathbb{R} and \mathbb{R}^3 : we say that a \mathbb{C} -valued function $Y(\theta, \phi)$ on the sphere is *square-integrable* if

$$(Y(\theta, \phi)|Y(\theta, \phi)) = \int_0^\pi \int_0^{2\pi} |Y(\theta, \phi)|^2 \sin \theta \, d\phi \, d\theta < \infty. \quad (5.19)$$

Just as with square-integrable functions on \mathbb{R} , the set of all such functions forms a Hilbert space, usually denoted as $L^2(S^2)$, where S^2 denotes the (two-dimensional) unit sphere in \mathbb{R}^3 . It is easy to see⁷ that each $\tilde{\mathcal{H}}_l \subset L^2(S^2)$, and in fact it turns out that all the $\tilde{\mathcal{H}}_l$ taken together are actually equal to $L^2(S^2)$! We will discuss this further in Sect. 5.6, but for now we note that this implies that the set $\{Y_m^l \mid 0 \leq l < \infty, -l \leq m \leq l\}$ of *all* the spherical harmonics form an (orthonormal) basis for $L^2(S^2)$. This can be thought of as a consequence of the *spectral theorems* of functional analysis, crucial to quantum mechanics, which tell us that, under suitable hypotheses, the eigenfunctions of a self-adjoint linear operator (in this case the spherical Laplacian Δ_{S^2}) form an orthonormal basis for the Hilbert space on which it acts. \square

⁶Just as in the previous example, however, one can define transformed coordinates θ' and ϕ' and then unitarity hinges on the Jacobian determinant $|\frac{\partial(\theta, \phi)}{\partial(\theta', \phi')}|$ being equal to one, which it is because rotations preserve area on the sphere.

⁷Each $Y \in \tilde{\mathcal{H}}_l$ is the restriction of a polynomial to S^2 and is hence continuous, hence $|Y|^2$ must have a finite maximum $M \in \mathbb{R}$. This implies

$$(Y|Y) \leq 4\pi M < \infty. \quad (5.20)$$

Exercise 5.7 Let $f \in \mathcal{H}_l(\mathbb{R}^3)$. Convince yourself that $\Pi_R f$ is also a degree l polynomial, and then use the chain rule to show that it is harmonic, hence an element of $\mathcal{H}_l(\mathbb{R}^3)$. The orthogonality of R should be crucial in your calculation!

Exercise 5.8 If you have never done so, find the induced $\mathfrak{so}(3)$ representation $\tilde{\pi}_l$ on $\tilde{\mathcal{H}}_l$ by expressing (5.15) in spherical coordinates. You should get

$$\begin{aligned}\tilde{\pi}_l(L_x) &= \sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \\ \tilde{\pi}_l(L_y) &= -\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \\ \tilde{\pi}_l(L_z) &= -\frac{\partial}{\partial \phi}.\end{aligned}$$

Check directly that these satisfy the $\mathfrak{so}(3)$ commutation relations, as they should.

The representations we have discussed so far have primarily been representations of matrix Lie groups and their associated Lie algebras. As we mentioned earlier, though, there are abstract Lie algebras which have physically relevant representations too. We will meet a couple of those now.

Example 5.10 *The Heisenberg algebra acting on $L^2(\mathbb{R})$*

We have mentioned this representation a few times already in this text, but we discuss it here to formalize it and place it in its proper context. The Heisenberg algebra $H = \text{Span}\{q, p, 1\} \subset \mathcal{C}(\mathbb{R}^2)$ has a unitary Lie algebra representation π on $L^2(\mathbb{R})$ given by

$$\begin{aligned}(\pi_q(f))(x) &= ix f(x) \\ (\pi_p(f))(x) &= -\frac{df}{dx}(x) \\ (\pi_1(f))(x) &= if(x).\end{aligned}\tag{5.21}$$

To verify that π is indeed a Lie algebra representation, one needs only to verify that the one non-trivial bracket is preserved, i.e. that $[\pi(q), \pi(p)] = \pi([q, p])$. This should be a familiar fact by now, and is readily verified if it is not. Showing that π is unitary requires a little bit of calculation, which you will perform below. The factors of i appearing above, especially the one in (5.21), may look funny; you should keep in mind, though, that the absence of these factors of i in the physics literature is again an artifact of the physicist's convention in defining Lie algebras, and that these factors of i are crucial for ensuring that the above operators are anti-Hermitian, as will be seen below. Note that the usual physics notation for these operators is $\hat{q} \equiv \pi_q = \pi(q)$ and $\hat{p} \equiv \pi_p = \pi(p)$.

Exercise 5.9 Verify that $\pi(q)$, $\pi(p)$, and $\pi(1)$ are all anti-Hermitian operators with respect to the usual inner product on $L^2(\mathbb{R})$. Exponentiate these operators to find, for $t, a, \theta \in \mathbb{R}$,

$$\begin{aligned}(e^{t\pi(q)} f)(x) &= e^{itx} f(x) \\ (e^{a\pi(p)} f)(x) &= f(x - a) \\ (e^{\theta\pi(1)} f)(x) &= e^{i\theta} f(x)\end{aligned}$$

and conclude, as we saw (in part) in Example 4.34, that $\pi(q)$ generates translations in momentum space, $\pi(p)$ generates translations in x , and $\pi(1)$ generates multiplication by a phase factor.

Example 5.11 *The adjoint representation of $\mathcal{C}(P)$*

Recall from Example 4.36 that for any Lie algebra \mathfrak{g} , regardless of whether or not it is the Lie algebra of a matrix Lie group, there is a Lie algebra homomorphism

$$\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$$

$$X \mapsto \text{ad}_X$$

and hence a representation of \mathfrak{g} on itself. Suppose that $\mathfrak{g} = \mathcal{C}(\mathbb{R}^6)$, the Lie algebra of observables on the phase space \mathbb{R}^6 which corresponds to a single particle living in three-dimensional space. What does the adjoint representation of $\mathcal{C}(\mathbb{R}^6)$ look like? Since $\mathcal{C}(\mathbb{R}^6)$ is infinite-dimensional, computing the matrix representations of basis elements is not really feasible. Instead, we pick a few important elements of $\mathcal{C}(\mathbb{R}^6)$ and determine how they act on $\mathcal{C}(\mathbb{R}^6)$ as linear differential operators. First, consider $\text{ad}_{q_i} \in \mathfrak{gl}(\mathcal{C}(\mathbb{R}^6))$. We have, for arbitrary $f \in \mathcal{C}(\mathbb{R}^6)$,

$$\text{ad}_{q_i} f = \{q_i, f\} = \frac{\partial f}{\partial p_i} \quad (5.22)$$

and so just as $-\frac{\partial}{\partial x}$ generates translation in the x -direction, $\text{ad}_{q_i} = \frac{\partial}{\partial p_i}$ generates translation in the p_i direction in phase space. Similarly, you can calculate that

$$\text{ad}_{p_i} = -\frac{\partial}{\partial q_i} \quad (5.23)$$

so that ad_{p_i} generates translation in the q_i direction. If $f \in \mathcal{C}(\mathbb{R}^6)$ depends only on the q_i and not the p_i , then one can show that

$$\text{ad}_{L_3} f = -\frac{\partial f}{\partial \phi} \quad (5.24)$$

where $\phi \equiv \tan^{-1}(q_2/q_1)$ is the azimuthal angle, so that L_3 generates rotations around the z -axis. Finally, for arbitrary $f \in \mathcal{C}(\mathbb{R}^6)$, one can show using Hamilton's equations⁸ that

$$\text{ad}_H = -\frac{d}{dt} \quad (5.25)$$

so that the Hamiltonian generates time translations. These facts are all part of the **Poisson Bracket formulation** of classical mechanics, and it is from this formalism that quantum mechanics gets the notion that the various 'symmetry generators' (which are of course just elements of the Lie algebra of the symmetry group G in question) that act on a Hilbert space should correspond to physical observables.

Exercise 5.10 Verify (5.23)–(5.25). If you did Exercise 4.30, you only need to verify (5.25), for which you will need Hamilton's equations, $\frac{\partial H}{\partial q_i} = \frac{dp_i}{dt}$, $\frac{\partial H}{\partial p_i} = \frac{dq_i}{dt}$.

⁸See Goldstein [6].

5.3 Tensor Product Representations

The next step in our study of representations is to learn how to take tensor products of representations. This is important for several reasons: First, as we will soon see, almost all representations of interest can be viewed as tensor products of other, more basic representations. Second, tensor products are ubiquitous in quantum mechanics (since they represent the addition of degrees of freedom), so we better know how they interact with representations. Finally, tensors can be understood as elements of tensor product spaces (cf. Sect. 3.5) and so understanding the tensor product of representations will allow us to understand more fully what is meant by the statement that a particular object ‘transforms like a tensor’.

Suppose, then, that we have two representations (Π_1, V_1) and (Π_2, V_2) of a group G . Then we can define their **tensor product representation** $(\Pi_1 \otimes \Pi_2, V_1 \otimes V_2)$ by

$$(\Pi_1 \otimes \Pi_2)(g) \equiv \Pi_1(g) \otimes \Pi_2(g) \in \mathcal{L}(V_1 \otimes V_2), \quad (5.26)$$

where you should recall from (3.69) how $\Pi_1(g) \otimes \Pi_2(g)$ acts on $V_1 \otimes V_2$. It is straightforward to check that this really does define a representation of G . If G is a matrix Lie group, we can then calculate the corresponding Lie algebra representation:

$$\begin{aligned} (\pi_1 \otimes \pi_2)(X) &\equiv \left. \frac{d}{dt} \Pi_1(e^{tX}) \otimes \Pi_2(e^{tX}) \right|_{t=0} \\ &= \lim_{h \rightarrow 0} \left[\frac{\Pi_1(e^{hX}) \otimes \Pi_2(e^{hX}) - I \otimes I}{h} \right] \\ &= \lim_{h \rightarrow 0} \left[\frac{\Pi_1(e^{hX}) \otimes \Pi_2(e^{hX}) - I \otimes \Pi_2(e^{hX}) + I \otimes \Pi_2(e^{hX}) - I \otimes I}{h} \right] \\ &= \lim_{h \rightarrow 0} \left[\frac{(\Pi_1(e^{hX}) - I)}{h} \otimes \Pi_2(e^{hX}) \right] + \lim_{h \rightarrow 0} \left[I \otimes \frac{(\Pi_2(e^{hX}) - I)}{h} \right] \\ &= \boxed{\pi_1(X) \otimes I + I \otimes \pi_2(X)} \end{aligned} \quad (5.27)$$

where in the second-to-last line we used the bilinearity of the tensor product. If we think of π_i as a sort of ‘derivative’ of Π_i , then one can think of (5.27) as a kind of product rule. In fact, the above calculation is totally analogous to the proof of the product rule from single-variable calculus! It is a nice exercise to directly verify that $\pi_1 \otimes \pi_2$ is a Lie algebra representation; this is Exercise 5.11 below.

Exercise 5.11 Verify that (5.27) defines a Lie algebra representation. Mainly, this consists of verifying that

$$[(\pi_1 \otimes \pi_2)(X), (\pi_1 \otimes \pi_2)(Y)] = (\pi_1 \otimes \pi_2)([X, Y]).$$

You may find the form of (5.27) familiar from the discussion below (3.69), as well as from other quantum mechanics texts; we are now in a position to explain this connection, as well as clarify what is meant by the terms “additive” and “multiplicative” quantum numbers.

Example 5.12 *Quantum mechanics, tensor product representations, and additive and multiplicative quantum numbers*

We have already discussed how in quantum mechanics one adds degrees of freedom by taking tensor products of Hilbert spaces. We have also discussed (in Example 4.33) how a matrix Lie group of symmetries of a physical system (i.e. a matrix Lie group G that acts on the phase space P and preserves the Hamiltonian H) gives rise to a Lie algebra of observables isomorphic to its own Lie algebra \mathfrak{g} , and how the quantum mechanical Hilbert space associated to that system should be a representation of \mathfrak{g} . Thus, if we have a composite physical system represented by a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, and if the \mathcal{H}_i carry representations π_i of some matrix Lie group of symmetries G , then it is natural to take as an additional axiom that G is represented on \mathcal{H} by the tensor product representation (5.26), which induces the representation (5.27) of \mathfrak{g} on \mathcal{H} .

For example, let G be the group of rotations $SO(3)$, and let $\mathcal{H}_1 = L^2(\mathbb{R}^3)$ correspond to the spatial degrees of freedom of a particle of spin s and $\mathcal{H}_2 = \mathbb{C}^{2s+1}$, $2s \in \mathbb{Z}$ correspond to the spin degree of freedom. Then $\mathfrak{so}(3)$ is represented on the total space $\mathcal{H} = L^2(\mathbb{R}^3) \otimes \mathbb{C}^{2s+1}$ by

$$(\pi \otimes \pi_s)(L_i) = \pi(L_i) \otimes I + I \otimes \pi_s(L_i) \quad (5.28)$$

where π is the representation of Example 5.8 and π_s is the spin s representation from Example 5.7. If we identify $(\pi \otimes \pi_s)(L_i)$ with \mathbf{J}_i , the i th component of the total angular momentum operator, and $\pi(L_i)$ with \mathbf{L}_i , the i th component of the orbital angular momentum operator, and $\pi_s(L_i)$ with \mathbf{S}_i , the i th component of the spin angular momentum operator, then (5.28) is just the component form of

$$\mathbf{J} = \mathbf{L} \otimes I + I \otimes \mathbf{S}, \quad (5.29)$$

the familiar equation expressing the total angular momentum as the sum of the spin and orbital angular momentum. We thus see that the form of this equation, which we were not in a position to understand (mathematically) when we first discussed it in Example 3.20, is dictated by representation theory, and in particular **by the form (5.27) of the induced representation of a Lie algebra on a tensor product space**. The same is true for other symmetry generators, like the translation generator p in the Heisenberg algebra. If we have two particles in one dimension with corresponding Hilbert spaces \mathcal{H}_i , $i = 1, 2$, along with representations π_i of the Heisenberg algebra, then the representation of p on the total space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ is just

$$(\pi_1 \otimes \pi_2)(p) = \pi_1(p) \otimes 1 + I \otimes \pi_2(p) \equiv \hat{p}_1 \otimes I + I \otimes \hat{p}_2 \quad (5.30)$$

where $\hat{p}_i \equiv \pi_i(p)$. This expresses the fact that the total momentum is just the sum of the momenta of the individual particles!

More generally, (5.27) can be seen as the mathematical expression of the fact that **physical observables corresponding to generators in the Lie algebra are additive**. More precisely, we have the following: let $v_i \in \mathcal{H}_i$, $i = 1, 2$ be eigenvectors of operators $\pi_i(A)$ with eigenvalues a_i , where A is an element of the Lie algebra of a symmetry group G . Then $v_1 \otimes v_2$ is an eigenvector of $(\pi_1 \otimes \pi_2)(A)$ with eigenvalue

$a_1 + a_2$. In other words, the eigenvalue a of A is an **additive quantum number**. Most familiar quantum numbers, such as energy, momentum and angular momentum, are additive quantum numbers, but there are exceptions. One such exception is the parity operator P . If we have two three-dimensional physical systems with corresponding Hilbert spaces \mathcal{H}_i , then the \mathcal{H}_i should furnish representations Π_i of $O(3)$. Now, $P = -I \in O(3)$ and if $v_i \in \mathcal{H}_i$ are eigenvectors of $\Pi_i(P)$ with eigenvalues λ_i , then $v_1 \otimes v_2 \in \mathcal{H}_1 \otimes \mathcal{H}_2$ has eigenvalue

$$\begin{aligned} ((\Pi_1 \otimes \Pi_2)(P))(v_1 \otimes v_2) &= \Pi_1(P)v_1 \otimes \Pi_2(P)v_2 \\ &= (\lambda_1 \lambda_2)v_1 \otimes v_2, \end{aligned}$$

where we used the property (3.49) of the tensor product in the last line. Thus parity is known as a **multiplicative quantum number**. This is due to the fact that the parity operator P is an element of the symmetry group, whereas most other observables are elements of the symmetry algebra (usually the Lie algebra corresponding to the symmetry group). \square

Our next order of business is to clarify what it means for an object to “transform like a tensor”. We already addressed this to a certain degree in Sect. 3.2 when we discussed change of bases. There, however, we looked at how a change of basis affects the component representation of tensors, which was the passive point of view. Here we will take the active point of view, where instead of changing bases we will be considering a group G acting on a vector space V via some representation Π . Taking the active point of view should nonetheless give the same transformation laws, and we will indeed see below that by considering representations of G on tensor products of V and V^* and examining their component form, we will reproduce the formulas from Sect. 3.2, just in the active form.

Recall from Sect. 3.5 that the set of tensors of rank (r, s) on a vector space V is just

$$T_s^r(V) = \underbrace{V^* \otimes \cdots \otimes V^*}_{r \text{ times}} \otimes \underbrace{V \otimes \cdots \otimes V}_{s \text{ times}}. \quad (5.31)$$

Given a representation Π of G on V , we would like to extend this representation to the vector space $T_s^r(V)$. To do this, we need to specify a representation of G on V^* . This is easily done: V^* is a vector space of functions on V (just the linear functions, in fact), so we can use (5.9) to obtain the **dual representation** (Π^*, V^*) , defined as

$$(\Pi_g^* f)(v) \equiv f(\Pi_{g^{-1}} v), \quad g \in G, \quad f \in V^*, \quad v \in V. \quad (5.32)$$

This representation has the nice property that if $\{e_i\}$ and $\{e^i\}$ are dual bases for V and V^* , then the bases $\{\Pi(g)e_i\}$ and $\{\Pi^*(g)e^i\}$ are also dual to each other for any $g \in G$. You will check this in Exercise 5.12 below.

With the dual representation in hand, we can then consider the tensor product representation $\Pi_s^r \equiv \underbrace{\Pi^* \otimes \cdots \otimes \Pi^*}_{r \text{ times}} \otimes \underbrace{\Pi \otimes \cdots \otimes \Pi}_{s \text{ times}}$, which is given by

$$\begin{aligned} & (\Pi_s^r(g))(f_1 \otimes \cdots \otimes f_r \otimes v_1 \otimes \cdots \otimes v_s) \\ &= \Pi^*(g)f_1 \otimes \cdots \otimes \Pi^*(g)f_r \otimes \Pi(g)v_1 \otimes \cdots \otimes \Pi(g)v_s. \end{aligned} \quad (5.33)$$

If G is a matrix Lie group, it is easy to check that the corresponding Lie algebra representation π_s^r is given by

$$\begin{aligned} & (\pi_s^r(g))(f_1 \otimes \cdots \otimes f_r \otimes v_1 \otimes \cdots \otimes v_s) \\ &= \pi^*(g)f_1 \otimes \cdots \otimes f_r \otimes v_1 \otimes \cdots \otimes v_s \\ &+ f_1 \otimes \pi^*(g)f_2 \otimes \cdots \otimes f_r \otimes v_1 \otimes \cdots \otimes v_s + \cdots \\ &+ f_1 \otimes \cdots \otimes \pi^*(g)f_r \otimes v_1 \otimes \cdots \otimes v_s \\ &+ f_1 \otimes \cdots \otimes f_r \otimes \pi(g)v_1 \otimes \cdots \otimes v_s \\ &+ f_1 \otimes \cdots \otimes f_r \otimes v_1 \otimes \pi(g)v_2 \otimes \cdots \otimes v_s + \cdots \\ &+ f_1 \otimes \cdots \otimes f_r \otimes v_1 \otimes \cdots \otimes \pi(g)v_s. \end{aligned} \quad (5.34)$$

We will get a handle on these formulas by considering several examples.

Exercise 5.12 Let (Π, V) be a representation of some group G and (Π^*, V^*) its dual representation. Show that if the bases $\{e_i\}$ and $\{e^i\}$ are dual to each other, then so are $\{\Pi(g)e_i\}$ and $\{\Pi^*(g)e^i\}$ for any $g \in G$.

Exercise 5.13 Alternatively, we could have defined Π_s^r by thinking of tensors as functions on V and using the idea behind (5.9) to get

$$\begin{aligned} & (\Pi_s^r(g)T)(v_1, \dots, v_r, f_1, \dots, f_s) \\ &= T(\Pi(g^{-1})v_1, \dots, \Pi(g^{-1})v_r, \Pi^*(g^{-1})f_1, \dots, \Pi^*(g^{-1})f_s). \end{aligned}$$

Expand T in components and show that this is equivalent to the definition (5.33).

Example 5.13 The dual representation

Let us consider (5.33) with $r = 1, s = 0$, in which case the vector space at hand is just V^* and our representation is just the dual representation (5.32). What does this representation look like in terms of matrices? For arbitrary $f \in V^*, v \in V$ we have

$$\begin{aligned} (\Pi_g^* f)(v) &= [\Pi_g^* f]^T[v] \\ &= ([\Pi_g^*][f])^T[v] \\ &= [f]^T[\Pi_g^*]^T[v] \end{aligned}$$

as well as

$$\begin{aligned} (\Pi_g^* f)(v) &= f(\Pi_{g^{-1}}v) \\ &= [f][\Pi_{g^{-1}}][v] \end{aligned}$$

which implies

$$[\Pi_g^*]^T = [\Pi_{g^{-1}}] = [\Pi_g^{-1}] = [\Pi_g]^{-1}.$$

(Why are those last two equalities true?) We thus obtain

$$\boxed{[\Pi_g^*] = [\Pi_g]^{-1T}}. \quad (5.35)$$

In other words, the matrices representing G on the dual space are just the inverse transposes of the matrices representing G on the original vector space! This is just the active transformation version of (3.30). Accordingly, if (Π, V) is the fundamental representation of $O(n)$ or $SO(n)$, then the matrices of the dual representation are identical to those of the fundamental, which is just another expression of the fact that dual vectors transform just like regular vectors under orthogonal transformations (passive *or* active).

If G is a matrix Lie group then (Π, V) induces a representation (π, V) of \mathfrak{g} , and hence (Π^*, V^*) should induce a representation (π^*, V^*) of \mathfrak{g} as well. What does this representation look like? Well, by the definition of π^* , we have (for any $X \in \mathfrak{g}$, $f \in V^*$, $v \in V$)

$$\begin{aligned} (\pi^*(X)f)(v) &= \left. \frac{d}{dt} (\Pi^*(e^{tX})f)(v) \right|_{t=0} \\ &= \left. \frac{d}{dt} f(\Pi(e^{-tX})v) \right|_{t=0} \\ &= f(-\pi(X)v). \end{aligned}$$

You will show below that in terms of matrices this means

$$\boxed{[\pi^*(X)] = -[\pi(X)]^T}. \quad (5.36)$$

Note again that if π is the fundamental representation of $O(n)$ or $SO(n)$ then $[\pi(X)] = X$ is antisymmetric and so the dual representation is identical to the original representation.

Exercise 5.14 Prove (5.36). This can be done a couple different ways, either by calculating the infinitesimal form of (5.35) (by letting $g = e^{tX}$ and differentiating at $t = 0$) or by a computation analogous to the derivation of (5.35).

Example 5.14 $(\Pi_1^1, \mathcal{L}(V))$: *The linear operator representation*

Now consider (5.33) with $(r, s) = (1, 1)$. Our vector space is then just $\mathcal{T}_1^1 = \mathcal{L}(V)$, the space of linear operators on V ! Thus, any representation of a group G on a vector space V leads naturally to a representation of G on the space of operators on V . What does this representation look like? We know from (5.33) that G acts on \mathcal{T}_1^1 by

$$\Pi_1^1(g)(f \otimes v) = (\Pi_g^* f) \otimes (\Pi_g v), \quad v \in V, f \in V^*, g \in G. \quad (5.37)$$

This is not very enlightening, though. To interpret this, consider $f \otimes v$ as a linear operator T on V , so that

$$T(w) \equiv f(w)v, \quad w \in V.$$

Then (careful working your way through these equalities!)

$$\begin{aligned}
 (\Pi_1^1(g)T)(w) &= ((\Pi_g^* f)(w))\Pi_g v \\
 &= f(\Pi_{g^{-1}} w)\Pi_g v \\
 &= \Pi_g(f(\Pi_{g^{-1}} w)v) \\
 &= \Pi_g(T(\Pi_g^{-1}(w))) \\
 &= (\Pi_g T \Pi_g^{-1})(w)
 \end{aligned}$$

so we have

$$\boxed{\Pi_1^1(g)T = \Pi_g T \Pi_g^{-1}}. \quad (5.38)$$

It is easy to check that this computation also holds for an arbitrary $T \in \mathcal{L}(V)$, since any such T can be written as a linear combination of terms of the form $f \otimes v$. Thus, (5.39) tell us that the tensor product representation of G on $V^* \otimes V = \mathcal{L}(V)$ is just the original representation acting on operators by similarity transformations! This should not be too surprising, and you perhaps could have guessed that this is how the action of G on V would extend to $\mathcal{L}(V)$. Representing (5.38) by matrices yields

$$[\Pi_1^1(g)T] = [\Pi_g][T][\Pi_g]^{-1} \quad (5.39)$$

which is just the active version of (3.33).

If G is a matrix Lie group we can also consider the induced Lie algebra rep $(\pi_1^1, \mathcal{L}(V))$, which according to (5.27) acts by

$$(\pi_1^1(X))(f \otimes v) = (\pi^*(X)f) \otimes v + f \otimes \pi(X)v. \quad (5.40)$$

Since we saw in Chap. 4 that the adjoint representation of G induces the adjoint representation of \mathfrak{g} , we expect that π_1^1 should act by the commutator. This is in fact the case, and you will show below that

$$\boxed{\pi_1^1(X)T = [\pi(X), T]}. \quad (5.41)$$

This, of course, reduces to a commutator of matrices when a basis is chosen. □

Exercise 5.15 Prove (5.41). As in Exercise 5.14, this can be done by either computing the infinitesimal form of (5.38), or performing a calculation similar to the one above (5.38), starting with the Lie algebra representation associated to (5.37).

It should come as no surprise that the examples above reproduced the formulas from Sect. 3.2, and in fact it is easy to show that the general tensor product representation (5.33) is just the active version of our tensor transformation law (3.22). Using the fact that

$$(\Pi_g^*)_i^j = (\Pi_{g^{-1}})_i^j, \quad (5.42)$$

which you will prove below, we have for an arbitrary (r, s) tensor T ,

$$\begin{aligned}
\Pi_s^r(g)T &= \Pi_s^r(g)(T_{i_1 \dots i_r}^{j_1 \dots j_s} e^{i_1} \otimes \dots \otimes e^{i_r} \otimes e_{j_1} \otimes \dots \otimes e_{j_s}) \\
&= T_{i_1 \dots i_r}^{j_1 \dots j_s} \Pi_g^* e^{i_1} \otimes \dots \otimes \Pi_g^* e^{i_r} \otimes \Pi_g e_{j_1} \otimes \dots \otimes \Pi_g e_{j_s} \\
&= (\Pi_g^{-1})_{k_1}^{i_1} \dots (\Pi_g^{-1})_{k_r}^{i_r} (\Pi_g)_{j_1}^{l_1} \dots (\Pi_g)_{j_s}^{l_s} T_{i_1 \dots i_r}^{j_1 \dots j_s} e^{k_1} \otimes \dots \otimes e^{k_r} \\
&\quad \otimes e_{l_1} \otimes \dots \otimes e_{l_s}
\end{aligned} \tag{5.43}$$

so that, relabeling dummy indices, we have

$$\begin{aligned}
&(\Pi_s^r(g)T)_{i_1 \dots i_r}^{j_1 \dots j_s} \\
&= (\Pi_g^{-1})_{i_1}^{k_1} \dots (\Pi_g^{-1})_{i_r}^{k_r} (\Pi_g)_{l_1}^{j_1} \dots (\Pi_g)_{l_s}^{j_s} T_{k_1 \dots k_r}^{l_1 \dots l_s}
\end{aligned} \tag{5.44}$$

which is just the active version of (3.22), with $\Pi(g)$ replacing A and $\Pi(g)^{-1}$ replacing A^{-1} .

Exercise 5.16 Verify (5.42).

5.4 Symmetric and Antisymmetric Tensor Product Representations

With the tensor product representation now in place, we can now consider symmetric and antisymmetric tensor product representations. Note that for tensors of type $(r, 0)$ or $(0, r)$ the tensor product representation (5.33) is symmetric, in the sense that all the factors in the tensor product are treated equally. A moment's thought then shows that if we have a completely symmetric tensor $T \in S^r(V)$, then $\Pi_r^0(g)T$ is also in $S^r(V)$. The same is true for the completely antisymmetric tensors $\Lambda^r V$, and for the spaces $S^r(V^*)$ and $\Lambda^r V^*$. Thus these subspaces of $\mathcal{T}_r^0(V)$ and $\mathcal{T}_0^r(V)$ furnish representations of G in their own right, and we will see below that many of these representations are familiar.

Example 5.15 $S^l(\mathbb{C}^2)$

Consider $S^l(\mathbb{C}^2)$, the completely symmetric $(0, l)$ tensors on \mathbb{C}^2 . By way of example, when $l = 3$ a basis for this space is given by

$$\begin{aligned}
v_0 &\equiv e_1 \otimes e_1 \otimes e_1 \\
v_1 &\equiv e_2 \otimes e_1 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_2 \\
v_2 &\equiv e_2 \otimes e_2 \otimes e_1 + e_1 \otimes e_2 \otimes e_2 + e_2 \otimes e_1 \otimes e_2 \\
v_3 &\equiv e_2 \otimes e_2 \otimes e_2
\end{aligned}$$

and in general we have

$$\begin{aligned}
v_0 &\equiv e_1 \otimes e_1 \otimes \dots \otimes e_1 \\
v_1 &\equiv e_2 \otimes e_1 \otimes \dots \otimes e_1 + \text{permutations} \\
v_2 &\equiv e_2 \otimes e_2 \otimes e_1 \otimes \dots \otimes e_1 + \text{permutations} \\
&\vdots
\end{aligned}$$

$$v_{l-1} \equiv e_2 \otimes e_2 \otimes \cdots \otimes e_1 + \text{permutations}$$

$$v_l \equiv e_2 \otimes e_2 \otimes \cdots \otimes e_2.$$

The tensor product representation $(\Pi_l^0, \mathcal{T}_l^0(\mathbb{C}^2))$ restricts to $S^l(\mathbb{C}^2) \subset \mathcal{T}_l^0(\mathbb{C}^2)$, yielding a representation of $SU(2)$ which we will denote as $(S^l\Pi, S^l(\mathbb{C}^2))$. One can easily compute the corresponding $\mathfrak{su}(2)$ representation, denoted $(S^l\pi, S^l(\mathbb{C}^2))$, and it is easy to show that in the basis given above,

$$[(S^l\pi)(S_z)] = i \begin{pmatrix} l/2 & & & \\ & l/2 - 1 & & \\ & & \ddots & \\ & & & -l/2 \end{pmatrix} \quad (5.45)$$

which is the same as (5.11). This suggests that $S^l(\mathbb{C}^2)$ is the same representation as $P_l(\mathbb{C}^2)$, and is thus also the same as the spin $l/2$ representation of $\mathfrak{su}(2)$. We will soon see that this is indeed the case. Note that we have a correspondence here between symmetric $(0, l)$ tensors on a vector space and degree l polynomials on a vector space, just as we did in Example 3.24.

Exercise 5.17 Verify (5.45).

Example 5.16 *Antisymmetric second rank tensors and the adjoint representation of $O(n)$*

Consider the fundamental representation (Π, \mathbb{R}^n) of $O(n)$. We can restrict the tensor product representation $(\Pi_2^0, \mathbb{R}^n \otimes \mathbb{R}^n)$ to the subspace $\Lambda^2\mathbb{R}^n$ to get the antisymmetric tensor product representation of $O(n)$ on $\Lambda^2\mathbb{R}^n$, which we will denote as $\Lambda^2\Pi$. Let $X = X^{ij}e_i \otimes e_j \in \Lambda^2\mathbb{R}^n$ with the standard basis. Then by (5.33), we have (for $R \in O(n)$)

$$\Lambda^2\Pi(R)(X) = X^{ij}Re_i \otimes Re_j = \sum_{k,l} X^{ij} R_{ki} R_{lj} e_k \otimes e_l \quad (5.46)$$

which in terms of matrices reads

$$[\Lambda^2\Pi(R)X] = R[X]R^T = R[X]R^{-1}. \quad (5.47)$$

So far we have just produced the active transformation law for a $(0, 2)$ tensor, and we have not yet made use of the fact that X is antisymmetric. Taking the antisymmetry of X into account, however, means that $[X]$ is an antisymmetric matrix, and (5.47) tells us that it transforms under $O(n)$ by similarity transformations. This, however, is an exact description of the adjoint representation of $O(n)$! So we conclude that the adjoint representation of $O(n)$ (and hence of $SO(n)$ and $\mathfrak{so}(n)$) are the same as the tensor product representation $(\Lambda^2\Pi, \Lambda^2\mathbb{R}^n)$. \square

Example 5.17 *Antisymmetric tensor representations of $O(3)$*

Consider the antisymmetric tensor representations $(\Lambda^k\Pi, \Lambda^k\mathbb{R}^3)$, $k = 1, \dots, 3$ of $O(3)$, obtained by restricting $(\Pi_k^0, \mathcal{T}_k^0(\mathbb{R}^3))$ to $\Lambda^k\mathbb{R}^3 \subset \mathcal{T}_k^0(\mathbb{R}^3)$. For convenience

we will define $\Lambda^0 \mathbb{R}^3$ to be the trivial representation on \mathbb{R} (also known as the *scalar* representation). Now, we already know that $\Lambda^1 \mathbb{R}^3 = \mathbb{R}^3$ is the fundamental representation, and from the previous example we know that $\Lambda^2 \mathbb{R}^3$ is the adjoint representation, also known as the pseudovector representation. What about $\Lambda^3 \mathbb{R}^3$? We know that this vector space is one-dimensional (why?), so is it just the trivial representation? Not quite. Taking the Levi-Civita tensor $e_1 \wedge e_2 \wedge e_3$ as our basis vector, we have

$$\begin{aligned} (\Lambda^3 \Pi(R))(e_1 \wedge e_2 \wedge e_3) &= (Re_1) \wedge (Re_2) \wedge (Re_3) \\ &= |R|e_1 \wedge e_2 \wedge e_3 \quad \text{by (3.105)} \\ &= \pm e_1 \wedge e_2 \wedge e_3. \end{aligned}$$

Thus $e_1 \wedge e_2 \wedge e_3$ is invariant under *rotations* but is still not a scalar, since it changes sign under inversion. An object that transforms this way is known as a *pseudoscalar*, and $\Lambda^3 \mathbb{R}^3$ is thus known as the pseudoscalar representation of $O(3)$. Note that if we restrict to $SO(3)$, then $|R| = 1$ and $\Lambda^3 \mathbb{R}^3$ is then just the scalar (trivial) representation, just as $\Lambda^2 \mathbb{R}^3$ is just the vector representation of $SO(3)$. Also, as we mentioned before, if we restrict our representations to $\mathbb{Z}_2 \simeq \{I, -I\} \subset O(3)$ then they reduce to either the trivial or the alternating representation. We summarize all this in the table below.

V	$O(3)$	$SO(3)$	$\{I, -I\} \simeq \mathbb{Z}_2$
$\Lambda^0 \mathbb{R}^3$	scalar	scalar	trivial
$\Lambda^1 \mathbb{R}^3$	vector	vector	alt
$\Lambda^2 \mathbb{R}^3$	pseudovector	vector	trivial
$\Lambda^3 \mathbb{R}^3$	pseudoscalar	scalar	alt

Example 5.18 *Antisymmetric tensor representations of $O(3, 1)$*

Here we repeat the analysis from the previous example but in the case of $O(3, 1)$. As above, $\Lambda^0 \mathbb{R}^4 \equiv \mathbb{R}$ is the trivial (scalar) representation, $\Lambda^1 \mathbb{R}^4 = \mathbb{R}^4$ is the vector representation, and $\Lambda^2 \mathbb{R}^4$ is just known as the second rank antisymmetric tensor representation (or sometimes just *tensor* representation). How about $\Lambda^3 \mathbb{R}^4$? To get a handle on that, we will compute matrix representations for the corresponding Lie algebra representation, using the following basis \mathcal{B} for $\Lambda^3 \mathbb{R}^4$:

$$\begin{aligned} f_1 &\equiv e_2 \wedge e_3 \wedge e_4 \\ f_2 &\equiv -e_1 \wedge e_3 \wedge e_4 \\ f_3 &\equiv e_1 \wedge e_2 \wedge e_4 \\ f_4 &\equiv e_1 \wedge e_2 \wedge e_3. \end{aligned}$$

You will check below that in this basis, the operators $(\Lambda^3 \pi)(\tilde{L}_i)$ and $(\Lambda^3 \pi)(K_i)$ are given by

$$[(\Lambda^3 \pi)(\tilde{L}_i)]_{\mathcal{B}} = \tilde{L}_i \quad (5.48)$$

$$[(\Lambda^3 \pi)(K_i)]_{\mathcal{B}} = K_i \quad (5.49)$$

so $(\Lambda^3 \pi, \Lambda^3 \mathbb{R}^4)$ is the same as the fundamental (vector) representation of $\mathfrak{so}(3, 1)$! When we consider the group representation $\Lambda^3 \Pi$, though, there is a slight difference between $\Lambda^3 \mathbb{R}^4$ and the vector representation; in the vector representation, the parity operator takes the usual form

$$\Pi(P) = P = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.50)$$

whereas on $\Lambda^3 \mathbb{R}^4$, we have (as you will check below)

$$[\Lambda^3 \Pi(P)]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (5.51)$$

which is equal to $-P$. Thus the elements of $\Lambda^3 \mathbb{R}^4$ transform like 4-vectors under infinitesimal Lorentz transformations (and, as we will show, under *proper* Lorentz transformations), but they transform with the wrong sign under parity, and so in analogy to the three-dimensional Euclidean case, $\Lambda^3 \mathbb{R}^4$ is known as the **pseudovector** representation of $O(3, 1)$. As in the Euclidean case, if one restricts to $SO(3, 1)_o$ then parity is excluded and then $\Lambda^3 \mathbb{R}^4$ and \mathbb{R}^4 are identical, but only as representations of the *proper* Lorentz group $SO(3, 1)_o$.

The next representation, $\Lambda^4 \mathbb{R}^4$, is one-dimensional but, as in the previous example, is not quite the trivial representation. As above, we compute the action of $\Lambda^4 \Pi(A)$, $A \in O(3, 1)$ on the Levi-Civita tensor $e_1 \wedge e_2 \wedge e_3 \wedge e_4$:

$$\begin{aligned} (\Lambda^4 \Pi(A))(e_1 \wedge e_2 \wedge e_3 \wedge e_4) &= (Ae_1) \wedge (Ae_2) \wedge (Ae_3) \wedge (Ae_4) \\ &= |A|e_1 \wedge e_2 \wedge e_3 \wedge e_4 \\ &= \pm e_1 \wedge e_2 \wedge e_3 \wedge e_4 \end{aligned}$$

and so $e_1 \wedge e_2 \wedge e_3 \wedge e_4$ is invariant under proper Lorentz transformations but changes sign under improper Lorentz transformations. As in the Euclidean case, such an object is known as a **pseudoscalar**, and so $(\Lambda^4 \Pi, \Lambda^4 \mathbb{R}^4)$ is known as the pseudoscalar representation of $O(3, 1)$.

As in the previous example, we can use $\{I, P\} \simeq \mathbb{Z}_2$ to distinguish between $\Lambda^3 \mathbb{R}^4$ and \mathbb{R}^4 , as well as between $\Lambda^4 \mathbb{R}^4$ and the trivial representation. The operators $\Pi(P)$ and $\Lambda^3 \Pi(P)$ can be distinguished in a basis-independent way by noting that $\Pi(P)$ is diagonalizable with eigenvalues $\{-1, -1, -1, 1\}$, whereas $\Lambda^3 \Pi(P)$ has eigenvalues $\{1, 1, 1, -1\}$. We again summarize in a table, where in the last column we write the eigenvalues of $\Lambda^k \Pi(P)$ in those cases where $\Lambda^k \Pi(P) \neq \pm I$:

V	$O(3, 1)$	$SO(3, 1)_o$	$\{I, P\} \simeq \mathbb{Z}_2$
$\Lambda^0 \mathbb{R}^4$	scalar	scalar	trivial
$\Lambda^1 \mathbb{R}^4$	vector	vector	$\{-1, -1, -1, 1\}$
$\Lambda^2 \mathbb{R}^4$	tensor	tensor	$\{1, 1, 1, -1, -1, -1\}$
$\Lambda^3 \mathbb{R}^4$	pseudovector	vector	$\{1, 1, 1, -1\}$
$\Lambda^4 \mathbb{R}^4$	pseudoscalar	scalar	alt

Exercise 5.18 By explicit computation, verify (5.48), (5.49), and (5.51). Also, show that $\Lambda^2 \Pi(P)$ has eigenvalues $\{-1, -1, -1, 1, 1, 1\}$. You may need to choose a basis for $\Lambda^2 \mathbb{R}^4$ to do this.

5.5 Equivalence of Representations

In the previous section we noted that the vector and dual vector representations of $O(n)$ (and hence $SO(n)$ and $\mathfrak{so}(n)$) were ‘the same’, as were the adjoint representation of $O(n)$ and the antisymmetric tensor product representation $(\Lambda^2 \Pi, \Lambda^2 \mathbb{R}^n)$. We had also noted a few equivalences in Sect. 5.1, where we pointed out that the adjoint representation of $SO(3)$ was, in a certain matrix representation, identical to the vector representation of $SO(3)$, and where we claimed that the adjoint representation of $\mathfrak{so}(3, 1)$ is equivalent to the antisymmetric second rank tensor representation. However, we never made precise what we meant when we said that two representations were ‘the same’ or ‘equivalent’, and in the cases where we attempted to prove such a claim, we usually just showed that the matrices of two representations were identical when particular bases were chosen. Defining equivalence in such a way is adequate but somewhat undesirable, as it requires a choice of basis; we would like an alternative definition that is more intrinsic and conceptual and that does not require a choice of coordinates. The desired definition is as follows: Suppose we have two representations (Π_1, V_1) and (Π_2, V_2) of a group G . A linear map $\phi : V_1 \rightarrow V_2$ which satisfies

$$\Pi_2(g)(\phi(v)) = \phi(\Pi_1(g)v) \quad \forall v \in V_1, g \in G \quad (5.52)$$

is said to be an *intertwining map* or *intertwiner*. This just means that the action of G via the representations commutes with the action of ϕ . If in addition ϕ is a vector space isomorphism, then (Π_1, V_1) and (Π_2, V_2) are said to be *equivalent*. Occasionally we will denote equivalence by $(\Pi_1, V_1) \simeq (\Pi_2, V_2)$. Equivalence of Lie algebra representations and their corresponding intertwining maps are defined similarly, by the equation

$$\pi_2(X)(\phi(v)) = \phi(\pi_1(X)v) \quad \forall v \in V_1, X \in \mathfrak{g}. \quad (5.53)$$

Another way to write (5.52) is as an equality between maps that go from V_1 to V_2 :

$$\Pi_2(g) \circ \phi = \phi \circ \Pi_1(g) \quad \forall g \in G \quad (5.54)$$

and likewise for Lie algebra representations. When ϕ is an isomorphism, this can be interpreted as saying that $\Pi_1(g)$ and $\Pi_2(g)$ are the ‘same’ map, once we use the intertwiner ϕ to identify V_1 and V_2 . Another way to interpret this is to choose bases for V_1 and V_2 and then write (5.54) as

$$[\Pi_2(g)][\phi] = [\phi][\Pi_1(g)] \quad \forall g \in G$$

or

$$[\Pi_2(g)] = [\phi][\Pi_1(g)][\phi]^{-1} \quad \forall g \in G$$

which says that the matrices $[\Pi_2(g)]$ and $[\Pi_1(g)]$ are related by a similarity transformation. You will use this below to show that our definition of equivalence of representations is equivalent to the statement that there exists bases for V_1 and V_2 such that $[\Pi_1(g)] = [\Pi_2(g)] \quad \forall g \in G$ (or the analogous statement for Lie algebras).

Exercise 5.19 Show that two representations (Π_1, V_1) and (Π_2, V_2) are equivalent if and only if there exist bases $\mathcal{B}_1 \subset V_1$ and $\mathcal{B}_2 \subset V_2$ such that

$$[\Pi_1(g)]_{\mathcal{B}_1} = [\Pi_2(g)]_{\mathcal{B}_2} \quad \forall g \in G. \quad (5.55)$$

Exercise 5.20 Let (Π_i, V_i) , $i = 1, 2$ be two equivalent representations of a group G , and let $H \subset G$ be a subgroup. Prove that restricting $\Pi_i : G \rightarrow GL(V_i)$ to maps $\Pi_i : H \rightarrow GL(V_i)$ yield representations of H , and that these representations of H are also equivalent. Thus, for example, equivalent representations of $O(n)$ yield equivalent representations of $SO(n)$, as one would expect.

Before we get to some examples, there are some immediate questions that arise. For instance, do equivalent representations of a matrix Lie group G give rise to equivalent representations of \mathfrak{g} , and conversely, do equivalent representations of \mathfrak{g} come from equivalent representations of G ? As to the first question, on philosophical grounds alone we would expect that since \mathfrak{g} consists of ‘infinitesimal’ group elements, equivalent group representations should yield equivalent Lie algebra representations. This is in fact the case:

Proposition 5.1 *Let G be a matrix Lie group, and let (Π_i, V_i) , $i = 1, 2$ be two equivalent representations of G with intertwining map ϕ . Then ϕ is also an intertwiner for the induced Lie algebra representations (π_i, V_i) , and so the induced Lie algebra representations are equivalent as well.*

Proof We proceed by direct calculation. Since ϕ is an intertwiner between (Π_1, V_1) and (Π_2, V_2) we have

$$\Pi_2(e^{tX})(\phi(v)) = \phi(\Pi_1(e^{tX})v) \quad \forall v \in V_1, X \in \mathfrak{g}, t \in \mathbb{R}.$$

Taking the t derivative of the above equation, evaluating at $t = 0$, and using the definition of the induced Lie algebra representations π_i , as well as the fact that ϕ commutes with derivatives (cf. Exercise 4.34), we obtain

$$\pi_2(X)(\phi(v)) = \phi(\pi_1(X)v) \quad \forall v \in V_1, X \in \mathfrak{g}$$

and so π_1 and π_2 are equivalent (you should explicitly confirm this as an exercise if more detail is needed). \square

The second question, of whether or not equivalent representations of \mathfrak{g} come from equivalent representations of G , is a bit trickier. After all, we noted in the introduction to this chapter (and will see very concretely in the case of $\mathfrak{so}(3)$) that not every representation of \mathfrak{g} necessarily comes from a representation of G . However, if we know that two equivalent Lie algebra representations (π_i, V_i) , $i = 1, 2$ actually do come from two group representations (Π_i, V_i) , and we know the group is connected, then it is true that Π_1 and Π_2 are equivalent.

Proposition 5.2 *Let G be a connected Lie group and let (Π_i, V_i) , $i = 1, 2$ be two representations of G , with associated Lie algebra representations (π_i, V_i) . If the Lie algebra representations (π_i, V_i) are equivalent, then so are the group representations (Π_i, V_i) from which they came.*

Proof The argument relies on the following fact, which we will not prove (see Hall [8] for details): if G is a *connected* matrix Lie group, then any $g \in G$ can be written as a product of exponentials. That is, for any $g \in G$ there exist $X_i \in \mathfrak{g}$, $t_i \in \mathbb{R}$, $i = 1, \dots, n$ such that

$$g = e^{t_1 X_1} e^{t_2 X_2} \dots e^{t_n X_n}. \quad (5.56)$$

(In fact, for all the connected matrix Lie groups we have met besides $SL(2, \mathbb{C})$, every group element can be written as a *single* exponential. For $SL(2, \mathbb{C})$, the polar decomposition theorem [see Problem 4.6] guarantees that any group element can be written as a product of *two* exponentials.) With this fact in hand we can show that Π_1 and Π_2 are equivalent. Let ϕ be an intertwining map between (π_1, V_1) and (π_2, V_2) . Then for any $g \in G$ and $v \in V_1$ we have

$$\begin{aligned} \Pi_2(g)(\phi(v)) &= \Pi_2(e^{t_1 X_1} e^{t_2 X_2} \dots e^{t_n X_n})(\phi(v)) \\ &= \Pi_2(e^{t_1 X_1}) \Pi_2(e^{t_2 X_2}) \dots \Pi_2(e^{t_n X_n})(\phi(v)) \quad \text{since } \Pi_2 \text{ is a homomorphism} \\ &= (e^{t_1 \pi_2(X_1)} e^{t_2 \pi_2(X_2)} \dots e^{t_n \pi_2(X_n)})(\phi(v)) \quad \text{by definition of } \pi_2 \\ &= \phi(e^{t_1 \pi_1(X_1)} e^{t_2 \pi_1(X_2)} \dots e^{t_n \pi_1(X_n)} v) \quad \text{by Exercise 5.21 below} \\ &= \phi(\Pi_1(e^{t_1 X_1}) \Pi_1(e^{t_2 X_2}) \dots \Pi_1(e^{t_n X_n}) v) \quad \text{by definition of } \pi_1 \\ &= \phi(\Pi_1(e^{t_1 X_1} e^{t_2 X_2} \dots e^{t_n X_n}) v) \quad \text{since } \Pi_1 \text{ is a homomorphism} \\ &= \phi(\Pi_1(g)v) \end{aligned}$$

and so Π_1 and Π_2 are equivalent. \square

Proposition 5.2 is useful in that it allows us to prove equivalence of group representations by examining the associated Lie algebra representations, which are (by virtue of linearity) often easier to work with.

Exercise 5.21 Let π_1 and π_2 be two equivalent representations of a Lie algebra \mathfrak{g} with intertwining map ϕ . Prove by expanding the exponential in a power series that

$$e^{t\pi_2(X)} \circ \phi = \phi \circ e^{t\pi_1(X)} \quad \forall X \in \mathfrak{g}.$$

Exercise 5.22 We claimed above that when a matrix Lie group G is connected, then any group element can be written as a product of exponentials as in (5.56). To see why the hypothesis of connectedness is important, consider the disconnected matrix Lie group $O(3)$ and find an element of $O(3)$ that *cannot* be written as a product of exponentials.

Now it is time for some examples.

Example 5.19 *Equivalence of $\mathfrak{so}(3)$ and \mathbb{R}^3 as $SO(3)$ representations*

We know that the adjoint representation and the fundamental representation of $\mathfrak{so}(3)$ are equivalent, by Exercise 5.19 and the fact that in the $\{L_i\}_{i=1-3}$ basis we have $[\text{ad}_{L_i}] = L_i$. What, then, is the intertwining map between \mathbb{R}^3 and $\mathfrak{so}(3)$? Simply the map

$$\begin{aligned} \phi : \mathfrak{so}(3) &\rightarrow \mathbb{R}^3 \\ \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} &\mapsto (x, y, z) \end{aligned}$$

which is just the map $X \mapsto [X]_{\mathcal{B}}$, where $\mathcal{B} = \{L_i\}_{i=1-3}$. To verify that this is an intertwiner we will actually work on the Lie algebra level, and then use Proposition 5.2 to conclude that the representations are equivalent on the group level. To verify that ϕ satisfies (5.53), one need only prove that the equation holds for an arbitrary *basis* element of $\mathfrak{so}(3)$; since the π_i are linear maps, we can expand any $X \in \mathfrak{so}(3)$ in terms of our basis and the calculation will reduce to verifying the equality just for the basis elements. (This is the advantage of working with Lie algebras; facts about representations [such as equivalence] are usually much easier to establish directly for Lie algebras than for the corresponding groups, since we can use linearity.) We thus calculate, for any $Y \in \mathfrak{so}(3)$,

$$\begin{aligned} (\phi \circ \text{ad}_{L_i})(Y) &= \phi(\text{ad}_{L_i} Y) \\ &= [\text{ad}_{L_i} Y] \\ &= [\text{ad}_{L_i}][Y] \\ &= L_i[Y] \end{aligned}$$

while

$$(L_i \circ \phi)(Y) = L_i[Y].$$

Thus ϕ is an intertwining map and the fundamental and adjoint representations of $\mathfrak{so}(3)$ are equivalent.

By Proposition 5.2, we can then conclude that the adjoint and fundamental representations of $SO(3)$ are equivalent, since $SO(3)$ is connected. What about the adjoint and fundamental representations of $O(3)$? You will recall that $O(3)$ is *not* connected

(and in fact has two separate connected components), so we cannot conclude that its adjoint and fundamental representations are equivalent. In fact, as we mentioned before, we know that these representations are *not* equivalent, since if $\phi : \mathfrak{so}(3) \rightarrow \mathbb{R}^3$ were an intertwining map, we would have for any $Y \in \mathfrak{so}(3)$,

$$\phi(\mathrm{Ad}_{-I} Y) = \phi(Y) \quad \text{since } \mathrm{Ad}_{-I} \text{ is the identity}$$

as well as

$$\begin{aligned} \phi(\mathrm{Ad}_{-I} Y) &= (-I)\phi(Y) \quad \text{since } \phi \text{ is an intertwiner} \\ &= -\phi(Y), \end{aligned}$$

a contradiction. Thus ϕ cannot exist, and the fundamental and adjoint representations of $O(3)$ are inequivalent.

Exercise 5.23 Use Example 5.17 and Exercise 5.19 to deduce that \mathbb{R} and $\Lambda^3 \mathbb{R}^3$ are equivalent as $SO(3)$ representations (in fact, they are both the trivial representation). Can you find an intertwiner? Do the same for the $SO(3, 1)_o$ representations \mathbb{R} and $\Lambda^4 \mathbb{R}^4$, using the results of Example 5.18. Using Proposition 5.2, also conclude that \mathbb{R}^4 and $\Lambda^3 \mathbb{R}^4$ are equivalent as $SO(3, 1)_o$ representations.

Exercise 5.24 Reread Example 3.29 in light of the last few sections. What would we now call the map J that we introduced in that example?

Example 5.20 Vector spaces with metrics and their duals

Let V be a vector space equipped with a metric g (recall that a metric is any symmetric, non-degenerate bilinear form), and let (Π, V) be a representation of G whereby G acts by isometries, i.e. $\Pi(g) \in \mathrm{Isom}(V) \quad \forall g \in G$. Examples of this include the fundamental representation of $O(n)$ on \mathbb{R}^n equipped with the Euclidean metric, or the fundamental representation of $O(n-1, 1)$ on \mathbb{R}^n with the Minkowski metric, but *not* $U(n)$ or $SU(n)$ acting on \mathbb{C}^n with the standard Hermitian inner product (why not?). If the assumptions above are satisfied, then (Π, V) is equivalent to the dual representation (Π^*, V^*) and the intertwiner is nothing but our old friend

$$\begin{aligned} L : V &\rightarrow V^* \\ v &\mapsto g(v, \cdot). \end{aligned}$$

You will verify in Exercise 5.25 that L is indeed an intertwiner, proving the asserted equivalence. Thus, in particular we again reproduce (this time in a basis-independent way) the familiar fact that dual vectors on n -dimensional Euclidean space transform just like ordinary vectors under orthogonal transformations. Additionally, we see that dual vectors on n -dimensional *Minkowski* space transform just like ordinary vectors under Lorentz transformations! See Exercise 5.26 for the matrix manifestation of this.

The reason we have excluded complex vector spaces with Hermitian inner products from this example is that in such circumstances, the map L is not linear (why not?) and thus cannot be an intertwiner. In fact, the fundamental representation of

$SU(n)$ on \mathbb{C}^n for $n \geq 3$ is *not* equivalent to its dual,⁹ and the dual representation in such circumstances can sometimes be interpreted as an *antiparticle* if the original representation represents a particle. For instance, a quark can be thought of as a vector in the fundamental representation \mathbb{C}^3 of $SU(3)$ (this is sometimes denoted as **3** in the physics literature), and then the dual representation \mathbb{C}^{3*} corresponds to an antiquark (this representation is often denoted as $\bar{\mathbf{3}}$). For $SU(2)$, however, the fundamental (spinor) representation *is* equivalent to its dual, though that does not follow from the discussion above. This equivalence is the subject of the next example.

Exercise 5.25 Let V be a vector space equipped with a metric g and let (Π, V) be a representation of a group G by isometries. Consider $L : V \rightarrow V^*$ as defined above. Prove that $L \circ \Pi_g = \Pi_g^* \circ L$ for all $g \in G$. Since $(L \circ \Pi_g)(v) \in V^*$ for any $v \in V$, you must show that

$$(L \circ \Pi_g)(v) = (\Pi_g^* \circ L)(v)$$

as *dual* vectors on V , which means showing that they have the same action on an arbitrary second vector w .

Exercise 5.26 Let (Π, \mathbb{R}^4) be the fundamental representation of $O(3, 1)$ on \mathbb{R}^4 , and let $\mathcal{B} = \{e_i\}_{i=1-4}$ be the standard basis for \mathbb{R}^4 and $\mathcal{B}^* = \{e^i\}_{i=1-4}$ the corresponding dual basis. Find another basis \mathcal{B}' for \mathbb{R}^{4*} such that

$$[\Pi^*(g)]_{\mathcal{B}'} = [\Pi(g)]_{\mathcal{B}}.$$

The map L might help you here. Does this generalize to the fundamental representation of $O(n-1, 1)$ on n -dimensional Minkowski space?

Example 5.21 *The fundamental (spinor) representation of $SU(2)$ and its dual*

Consider the fundamental representation (Π, \mathbb{C}^2) of $SU(2)$ and its dual representation (Π^*, \mathbb{C}^{2*}) . We will show that these representations are equivalent, by first showing that the induced Lie algebra representations are equivalent and then invoking Proposition 5.2. We will show equivalence of the Lie algebra representations by exhibiting a basis for \mathbb{C}^{2*} that yields the same matrix representations for (π^*, \mathbb{C}^{2*}) as for (π, \mathbb{C}^2) . For an intrinsic (coordinate-independent) proof, see Problem 5.2.

As should be familiar by now, the fundamental representation of $\mathfrak{su}(2)$ is just the identity map:

$$\begin{aligned}\pi(S_x) &= S_x = \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ \pi(S_y) &= S_y = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ \pi(S_z) &= S_z = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.\end{aligned}$$

This combined with (5.36) tells us that, in the standard dual basis \mathcal{B}^* ,

⁹We will not prove this here; see Hall [8] for details.

$$\begin{aligned} [\pi^*(S_x)] &= \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \\ [\pi^*(S_y)] &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ [\pi^*(S_z)] &= \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \end{aligned}$$

Now define a new basis $\mathcal{B}^{*'} = \{e^{1'}, e^{2'}\}$ for \mathbb{C}^{2*} by

$$\begin{aligned} e^{1'} &\equiv e^2 \\ e^{2'} &\equiv -e^1. \end{aligned}$$

You can check that the corresponding change of basis matrix A is

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (5.57)$$

and that in this basis the operators $\pi^*(S_i)$ are given by

$$\begin{aligned} [\pi^*(S_x)]_{\mathcal{B}^{*'}} &= A[\pi^*(S_x)]_{\mathcal{B}^*} A^{-1} = S_x \\ [\pi^*(S_y)]_{\mathcal{B}^{*'}} &= A[\pi^*(S_y)]_{\mathcal{B}^*} A^{-1} = S_y \\ [\pi^*(S_z)]_{\mathcal{B}^{*'}} &= A[\pi^*(S_z)]_{\mathcal{B}^*} A^{-1} = S_z \end{aligned} \quad (5.58)$$

and thus (π^*, \mathbb{C}^{2*}) is equivalent to (π, \mathbb{C}^2) . The connectedness of $SU(2)$ and Proposition 5.2 then imply that (Π^*, \mathbb{C}^{2*}) is equivalent to (Π, \mathbb{C}^2) , as claimed.

Exercise 5.27 Verify (5.57) and (5.58).

Exercise 5.28 Extend the above argument to show that the fundamental representation of $SL(2, \mathbb{C})$ is equivalent to its dual.

Example 5.22 *Antisymmetric second rank tensors and the adjoint representation of orthogonal and Lorentz groups*

We claimed in Sect. 5.1 that for the Lorentz group, the adjoint representation is ‘the same’ as the antisymmetric second rank tensor representation (the latter being the representation to which the electromagnetic field tensor $F^{\mu\nu}$ belongs). We also argued at the end of Sect. 5.3 that the same is true for the orthogonal group $O(n)$. Now it is time to precisely and rigorously prove these claims. Consider \mathbb{R}^n equipped with a metric g , where g is either the Euclidean metric or the Minkowski metric. The isometry group $G = \text{Isom}(V)$ is then either $O(n)$ or $O(n-1, 1)$, and \mathfrak{g} is then either $\mathfrak{so}(n)$ or $\mathfrak{so}(n-1, 1)$. To prove equivalence, we need an intertwiner $\phi: \Lambda^2 \mathbb{R}^n \rightarrow \mathfrak{g}$. Here we will define ϕ abstractly and use coordinate-free language to prove that it is an intertwiner. The proof is a little long, requires some patience, and may be skipped on a first reading, but it is a good exercise for getting acquainted with the machinery of this chapter; for an illuminating coordinate proof, see Problem 5.1.

To define ϕ abstractly we interpret \mathfrak{g} not as a space of matrices but rather as linear operators on \mathbb{R}^n . Since $\Lambda^2 \mathbb{R}^n$ is just the set of antisymmetric $(0, 2)$ tensors on \mathbb{R}^n ,

we can then define ϕ by just using the map L to ‘lower an index’ on an element $T \in \Lambda^2 \mathbb{R}^n$, converting the $(0, 2)$ tensor into a $(1, 1)$ tensor, i.e. a linear operator. More precisely, we define the linear operator $\phi(T)$ as

$$\phi(T)(v, f) \equiv T(L(v), f), \quad v \in \mathbb{R}^n, f \in \mathbb{R}^{n*}.$$

We must check, though, that $\phi(T) \in \mathfrak{g}$. Earlier, we had characterized \mathfrak{g} by a matrix condition like antisymmetry or the condition (4.70). Now that we are thinking of \mathfrak{g} as a space of linear operators, however, we need another characterization. Fortunately, we have already discussed this; as in (5.1), we can write

$$g(e^{tX}v, e^{tX}w) = g(v, w) \quad \forall v, w \in V, X \in \mathfrak{g}, t \in \mathbb{R}$$

and then differentiate at $t = 0$ to obtain

$$g(Xv, w) + g(v, Xw) = 0, \quad (5.59)$$

an alternative definition of \mathfrak{g} which you can easily show is equivalent to the matrix conditions we used originally (in fact, just writing out (5.59) in matrices leads to the antisymmetry condition in the Euclidean case, and (4.70) in the Lorentzian case). Thus, to show that the range of ϕ really is \mathfrak{g} , we just need to show that $\phi(T)$ as defined above satisfies (5.59):

$$\begin{aligned} & g(\phi(T)v, w) + g(v, \phi(T)w) \\ &= \phi(T)(v, L(w)) + \phi(T)(w, L(v)) \quad \text{by definition of } L \\ &= T(L(v), L(w)) + T(L(w), L(v)) \quad \text{by definition of } \phi(T) \\ &= 0 \quad \text{by antisymmetry of } T. \end{aligned}$$

Thus $\phi(T)$ really is in \mathfrak{g} . To show that ϕ is an intertwiner, we need to show that $\phi \circ \Lambda^2 \Pi(R) = \text{Ad}(R) \circ \phi$ for all $R \in G$. To do this, we will employ the alternate definition of the tensor product representation given in Exercise 5.13, which in this case says

$$(\Lambda^2 \Pi(R)T)(f, h) = T(\Pi^*(R^{-1})f, \Pi^*(R^{-1})h), \quad R \in G, f, h \in \mathbb{R}^{n*}. \quad (5.60)$$

We then have, for all $v \in \mathbb{R}^n, f \in \mathbb{R}^{n*}$ (careful with all the parentheses!),

$$\begin{aligned} ((\phi \circ \Lambda^2 \Pi(R))(T))(v, f) &= (\Lambda^2 \Pi(R)(T))(L(v), f) \\ &= T(\Pi^*(R^{-1})L(v), \Pi^*(R^{-1})f) \\ &= T(L(R^{-1}v), \Pi^*(R^{-1})f) \end{aligned}$$

where in the last equality we used the fact that L is an intertwiner between \mathbb{R}^n and \mathbb{R}^{n*} . We also have

$$\begin{aligned} ((\text{Ad}(R) \circ \phi)(T))(v, f) &= \phi(T)(R^{-1}v, \Pi^*(R^{-1})f) \\ &= T(L(R^{-1}v), \Pi^*(R^{-1})f) \end{aligned}$$

and we can thus conclude that $\phi \circ (\Lambda^2 \Pi(R)) = \text{Ad}(R) \circ \phi$, as desired.

So what does all this tell us? The conclusion that the tensor product representation of G on antisymmetric second rank tensors coincides with the adjoint representation of G on \mathfrak{g} is not at all surprising in the Euclidean case, because there \mathfrak{g} is just $\mathfrak{so}(n)$, the set of all antisymmetric matrices! In the Lorentzian case, however, we might be a little surprised, since the matrices in $\mathfrak{so}(n-1, 1)$ are not all antisymmetric. These matrices, however, represent linear operators, and if we use L to convert them into $(0, 2)$ tensors (via ϕ^{-1}) then they *are* antisymmetric! In other words, **we can always think of the Lie algebra of a group of metric-preserving operators as being antisymmetric tensors**, though we may have to raise or lower an index (via L) to make this manifest.

This is actually quite easy to show in coordinates; using the standard basis for \mathbb{R}^n and the corresponding components of $X = X_i^j e^i \otimes e_j \in \mathfrak{g}$ (still viewed as a linear operator on \mathbb{R}^n), we can plug two basis vectors into (5.59) and obtain

$$\begin{aligned} 0 &= g(Xe_i, e_j) + g(e_i, Xe_j) \\ &= X_i^k g(e_k, e_j) + X_j^k g(e_i, e_k) \\ &= X_i^k g_{kj} + X_j^k g_{ik} \\ &= X_{ij} + X_{ji} \end{aligned}$$

and so the $(2, 0)$ tensor corresponding to $X \in \mathfrak{g}$ is antisymmetric! \square

Our last example takes the form of an exercise, in which you will show that some of the representations of $SU(2)$ on $P_1(\mathbb{C}^2)$, as described in Example 5.7, are equivalent to more basic $SU(2)$ representations.

Exercise 5.29 Prove (by exhibiting an intertwining map) that $(\pi_1, P_1(\mathbb{C}^2))$ from Example 5.7 is equivalent to the fundamental representation of $\mathfrak{su}(2)$. Conclude (since $SU(2)$ is connected) that $(\Pi_1, P_1(\mathbb{C}^2))$ is equivalent to the fundamental representation of $SU(2)$. Do the same for $(\pi_2, P_2(\mathbb{C}))$ and the adjoint representation of $\mathfrak{su}(2)$ (you will need to consider a new basis that consists of *complex* linear combinations of the elements of B_2).

5.6 Direct Sums and Irreducibility

One of our goals in this chapter is to organize the various representations we have met into a coherent scheme, and to see how they are all related. Defining a notion of equivalence was the first step, so that we would know when two representations are ‘the same’. With that in place, we would now like to determine all the (inequivalent) representations of a given group or Lie algebra. In general this is a difficult problem, but for most of the matrix Lie groups we have met so far and their associated Lie algebras, there is a very nice way to do this: for each group or algebra, there exists a denumerable set of inequivalent representations (known as the “irreducible” representations) out of which all other representations can be built. Once these irreducible representations are known, any other representation can be broken down into a kind of “sum” of its irreducible components. In this section we will present the notions

of irreducibility and sum of vector spaces, and in subsequent sections we will enumerate all the irreducible representations of $SU(2)$, $SO(3)$, $SO(3, 1)$, $SL(2, \mathbb{C})$ and their associated Lie algebras.

To motivate the discussion, consider the vector space $M_n(\mathbb{R})$ of all $n \times n$ real matrices. There is a representation Π of $O(n)$ on this vector space given by similarity transformations:

$$\Pi(R)A \equiv RAR^{-1}, \quad R \in O(n), \quad A \in M_n(\mathbb{R}).$$

If we consider $M_n(\mathbb{R})$ to be the matrices corresponding to elements of $\mathcal{L}(\mathbb{R}^n)$, then this is just the matrix version of the linear operator representation Π_1^1 described above, with $V = \mathbb{R}^n$. Alternatively, it can be viewed as the matrix version of the representation Π_2^0 on $\mathbb{R}^n \otimes \mathbb{R}^n$, with identical matrix transformation law

$$[\Pi_2^0(R)(T)] = R[T]R^T = R[T]R^{-1}, \quad T \in \mathbb{R}^n \otimes \mathbb{R}^n.$$

Now, it turns out that there are some special properties that $A \in M_n(\mathbb{R})$ could have that would be preserved by $\Pi(R)$. For instance, if A is symmetric or antisymmetric, then RAR^{-1} is also; you can check this directly, or see it as a corollary of the discussion above Example 5.16. Furthermore, if A has zero trace, then so does RAR^{-1} . In fact, we can decompose A into a symmetric piece and an antisymmetric piece, and then further decompose the symmetric piece into a traceless piece and a piece proportional to the identity, as follows:

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) \tag{5.61}$$

$$= \frac{1}{n}(\text{Tr } A)I + \frac{1}{2}\left(A + A^T - \frac{2}{n}(\text{Tr } A)I\right) + \frac{1}{2}(A - A^T). \tag{5.62}$$

(You should check explicitly that the first term in (5.62) is proportional to the identity, the second is symmetric and traceless, and the third is antisymmetric.) Furthermore, this decomposition is unique, as you will show below. If we recall the definitions

$$\begin{aligned} S_n(\mathbb{R}) &= \{M \in M_n(\mathbb{R}) \mid M = M^T\} \\ A_n(\mathbb{R}) &= \{M \in M_n(\mathbb{R}) \mid M = -M^T\}, \end{aligned}$$

and add the new definitions

$$\begin{aligned} S'_n(\mathbb{R}) &\equiv \{M \in M_n(\mathbb{R}) \mid M = M^T, \text{Tr } M = 0\} \\ \mathbb{R}I &\equiv \{M \in M_n(\mathbb{R}) \mid M = cI, c \in \mathbb{R}\}, \end{aligned}$$

then this means that any $A \in M_n(\mathbb{R})$ can be written uniquely as a sum of elements of S'_n , A_n and $\mathbb{R}I$, all of which are subspaces of $M_n(\mathbb{R})$. This type of situation turns up frequently, so we formalize it with the following definition: If V is a vector space with subspaces W_1, W_2, \dots, W_k such that any $v \in V$ can be written *uniquely* as $v = w_1 + w_2 + \dots + w_k$, where $w_i \in W_i$, then we say that V is the **direct sum** of the W_i and we write

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

which is sometimes abbreviated as $V = \bigoplus_{i=1}^k W_i$. The decomposition of a vector v is sometimes written as $v = (w_1, \dots, w_k)$. In the situation above we have, as you can check,

$$M_n(\mathbb{R}) = S_n(\mathbb{R}) \oplus A_n(\mathbb{R}) = \mathbb{R}I \oplus S'_n(\mathbb{R}) \oplus A_n(\mathbb{R}). \quad (5.63)$$

Exercise 5.30 Show that if $V = \bigoplus_{i=1}^k W_i$, then $W_i \cap W_j = \{0\} \forall i \neq j$, i.e. the intersection of two different W_i is just the zero vector. Verify this explicitly in the case of the two decompositions in (5.63).

Exercise 5.31 Show that $V = \bigoplus_{i=1}^k W_i$ is equivalent to the statement that the set

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k,$$

where each \mathcal{B}_i is an arbitrary basis for W_i , is a basis for V .

Exercise 5.32 Show that $M_n(\mathbb{C}) = H_n(\mathbb{C}) \oplus u(n)$.

Our discussion above shows that all the subspaces that appear in (5.63) are *invariant*, in the sense that the action of $\Pi(R)$ on any element of one of the subspaces produces another element of the same subspace (i.e. if A is symmetric then $\Pi(R)A$ is also, etc.). In fact, we can define an *invariant subspace* of a group representation (Π, V) as a subspace $W \subset V$ such that $\Pi(g)w \in W$ for all $g \in G$, $w \in W$. Invariant subspaces of Lie algebra representations are defined analogously. Notice that the entire vector space V , as well as the zero vector $\{0\}$, are always (trivially) invariant subspaces. An invariant subspace $W \subset V$ that is neither equal to V nor to $\{0\}$ is said to be a *non-trivial* invariant subspace. Notice also that the invariance of W under the $\Pi(g)$ means we can restrict each $\Pi(g)$ to W (that is, interpret each $\Pi(g)$ as an operator $\Pi(g)|_W \in GL(W)$) and so obtain a representation $(\Pi|_W, W)$ of G on W . Given a representation V , the symmetric and antisymmetric subspaces $S^r(V)$ and $\Lambda^r(V)$ of the tensor product representation $\mathcal{T}_r^0(V)$ are nice examples of non-trivial invariant subspaces (recall, of course, that $\Lambda^r(V)$ is only non-trivial when $r \leq \dim V$).

Non-trivial invariant subspaces are very important in representation theory, as they allow us to block diagonalize the matrices corresponding to our operators. If we have a representation (Π, V) of a group G where V decomposes into a direct sum of invariant subspaces W_i ,

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k,$$

and if \mathcal{B}_i are bases for W_i and we take the union of the \mathcal{B}_i as a basis for V , then the matrix representation of the operator $\Pi(g)$ in this basis will look like

$$[\Pi(g)]_{\mathcal{B}} = \begin{pmatrix} [\Pi(g)]_{\mathcal{B}_1} & & & \\ & [\Pi(g)]_{\mathcal{B}_2} & & \\ & & \ddots & \\ & & & [\Pi(g)]_{\mathcal{B}_k} \end{pmatrix} \quad (5.64)$$

where each $[\Pi(g)]_{\mathcal{B}_i}$ is the matrix of $\Pi(g)$ restricted to the subspace W_i .

Thus, given a finite-dimensional representation (Π, V) of a group or Lie algebra, we can try to get a handle on it by decomposing V into a direct sum of (two or more) invariant subspaces. Each of these invariant subspaces forms a representation in its own right, and we could then try to further decompose these representations, iterating until we get a decomposition $V = W_1 \oplus \cdots \oplus W_k$ in which each of the W_i have no non-trivial invariant subspaces (if they did, we might be able to decompose further). These elementary representations play an important role in the theory, as we will see, and so we give them a name: we say that a representation W that has no non-trivial invariant subspaces is an **irreducible** representation (or *irrep*, for short). Furthermore, if a representation (Π, V) admits a decomposition $V = W_1 \oplus \cdots \oplus W_k$ where each of the W_i are irreducible, then we say that (Π, V) is **completely reducible**.¹⁰ If the decomposition consists of two or more summands, then we say that V is **decomposable**.¹¹ Thus we have the following funny-sounding sentence: if (Π, V) is completely reducible, then it is either decomposable or irreducible!

You may be wondering at this point how a finite-dimensional representation could not be completely reducible; after all, it is either irreducible or it contains a non-trivial invariant subspace W ; cannot we then decompose V into W and some subspace W' complementary to W ? We can, but the potential problem is that W' may not be invariant; that is, the group or Lie algebra action might take vectors in W' to vectors that do not lie in W' . For an example of this, see Problem 5.7. However, there do exist many groups and Lie algebras for whom every finite-dimensional representation *is* completely reducible. Such groups and Lie algebras are said to be **semi-simple**.¹² (Most of the matrix Lie groups we have met and their associated Lie algebras are semi-simple, but some of the abstract Lie algebras we have seen [like the Heisenberg algebra], as well as the matrix Lie groups $U(n)$, are not.) Thus, an arbitrary finite-dimensional representation of a semi-simple group or Lie algebra can by definition always be written as a direct sum of irreducible representations. (This is what we did when we wrote $M_n(\mathbb{R})$ as $M_n(\mathbb{R}) = \mathbb{R}I \oplus S'_n(\mathbb{R}) \oplus A_n(\mathbb{R})$, though we cannot yet prove that the summands are irreducible.) If we know all the irreducible representations of a given semi-simple group or Lie algebra, we then have a complete classification of *all* the finite-dimensional representations of that group or algebra, since any representation decomposes into a finite sum of irreps. This makes the determination of irreps an important task, which we will complete in this chapter for our favorite Lie algebras.

Example 5.23 Decomposition of tensor product representations

Say we have a semi-simple group or Lie algebra and two irreps V_1 and V_2 . The tensor product representation $V_1 \otimes V_2$ is usually *not* irreducible, but since our group

¹⁰Note that an irreducible representation is, trivially, completely reducible, since $V = V$ is a decomposition into irreducibles. Thus ‘irreducible’ and ‘completely reducible’ are not mutually exclusive categories, even if they may sound like it!

¹¹This terminology is not standard but will prove useful.

¹²Semi-simplicity can be defined in a number of equivalent ways, all of which are important. For more, see Hall [8] or Varadarajan [17].

or Lie algebra is semi-simple we can decompose $V_1 \otimes V_2$ as

$$V_1 \otimes V_2 = W_1 \oplus \cdots \oplus W_k$$

where the W_i are irreducible. In fact, the last decomposition in (5.63) is just the matrix version of the $O(n)$ -invariant decomposition

$$\mathcal{T}_2^0(\mathbb{R}^n) = \mathbb{R}^n \otimes \mathbb{R}^n = \mathbb{R}g \oplus S^{2'}(\mathbb{R}^n) \oplus \Lambda^2(\mathbb{R}^n) \quad (5.65)$$

where

$$g = \sum_i e_i \otimes e_i \quad (5.66)$$

$$S^{2'}(\mathbb{R}^n) = \{T \in S^2(\mathbb{R}^n) \mid \delta_{ij} T^{ij} = 0\} \quad \text{where } \delta_{ij} \text{ is Kronecker delta.} \quad (5.67)$$

Another familiar instance of the decomposition of a tensor product representation into irreducibles is when we are dealing with $SU(2)$ representations. As discussed in Example 3.21, it is well known that if we have the spin j representation $\mathbb{C}^{2j+1} \equiv V_j$ and spin j' representation $\mathbb{C}^{2j'+1} \equiv V_{j'}$ (we will discuss these representations thoroughly in the next section), then their tensor product decomposes as

$$V_j \otimes V_{j'} = V_{|j-j'|} \oplus V_{|j-j'|+1} \oplus \cdots \oplus V_{j+j'}. \quad (5.68)$$

The process of explicitly constructing this decomposition is known in the physics literature as the “addition of angular momentum”.

Example 5.24 *Decomposition of $\Lambda^2 \mathbb{R}^4$ as an $O(3)$ representation*

Consider the second rank antisymmetric tensor representation of $O(3, 1)$ on $\Lambda^2 \mathbb{R}^4$, but restrict the representation to $O(3)$ where we view $O(3) \subset O(3, 1)$ in the obvious way. This representation of $O(3)$ is reducible, since clearly $\Lambda^2 \mathbb{R}^3 \subset \Lambda^2 \mathbb{R}^4$ is an $O(3)$ -invariant subspace spanned by

$$\begin{aligned} f_1 &\equiv e_2 \wedge e_3 \\ f_2 &\equiv e_3 \wedge e_1 \\ f_3 &\equiv e_1 \wedge e_2. \end{aligned} \quad (5.69)$$

There is also a complementary invariant subspace, spanned by

$$\begin{aligned} f_4 &\equiv e_1 \wedge e_4 \\ f_5 &\equiv e_2 \wedge e_4 \\ f_6 &\equiv e_3 \wedge e_4. \end{aligned} \quad (5.70)$$

This second subspace is clearly equivalent to the vector representation of $O(3)$ since $O(3)$ leaves e_4 unaffected (the unconvinced reader can quickly check that the map $\phi : e_i \wedge e_4 \mapsto e_i$ is an intertwiner). Thus, as an $O(3)$ representation, $\Lambda^2 \mathbb{R}^4$ decomposes into the vector and pseudovector representations, i.e.

$$\Lambda^2 \mathbb{R}^4 \simeq \mathbb{R}^3 \oplus \Lambda^2 \mathbb{R}^3. \quad (5.71)$$

We can interpret this physically in the case of the electromagnetic field tensor, which lives in $\Lambda^2\mathbb{R}^4$; in that case, (5.71) says that **under (proper and improper) rotations, some components of the field tensor transform amongst themselves as a vector, and others as a pseudovector**. The components that transform as a vector comprise the electric field, and those that transform like a pseudovector comprise the magnetic field. To see this explicitly, one can think about which basis vectors from (5.70) and (5.69) go with which components of the matrix in (3.62).

One could, of course, further restrict this representation to $SO(3) \subset O(3, 1)$; one would get the same decomposition (5.71), except that for $SO(3)$ the representations \mathbb{R}^3 and $\Lambda^2\mathbb{R}^3$ are equivalent, and so under (proper) rotations the field tensor transforms as a pair of vectors, the electric field “vector” and the magnetic field “vector”. Of course, how these objects really transform depends on what transformation group you are talking about! The electric field and magnetic field both transform as vectors under rotations, but under improper rotations the electric field transforms as a vector and the magnetic field as a pseudovector. Furthermore, under Lorentz transformations the electric and magnetic fields cannot be meaningfully distinguished, as they transform together as the components of an antisymmetric second rank tensor! Physically, this corresponds to the fact that boosts can turn electric fields in one reference frame into magnetic fields in another, and vice versa.

Example 5.25 *Decomposition of $L^2(S^2)$ into irreducibles*

A nice example of a direct sum decomposition of a representation into its irreducible components is furnished by $L^2(S^2)$. As we noted in Example 5.9, the spectral theorems of functional analysis tell us that the eigenfunctions Y_m^l of the spherical Laplacian Δ_{S^2} form an orthogonal basis for the Hilbert space $L^2(S^2)$. We already know, though, that the Y_m^l of fixed l form a basis for $\tilde{\mathcal{H}}_l$, the spherical harmonics of degree l . We will show in the next section that each of the $\tilde{\mathcal{H}}_l$ is an irreducible $SO(3)$ representation, so we can decompose $L^2(S^2)$ into irreducible representations as

$$L^2(S^2) = \bigoplus_l \tilde{\mathcal{H}}_l = \tilde{\mathcal{H}}_1 \oplus \tilde{\mathcal{H}}_2 \oplus \tilde{\mathcal{H}}_3 \oplus \cdots. \quad \square$$

Before concluding this section we should point out that the notion of direct sum is useful not only in decomposing a given vector space into mutually exclusive subspaces, but also in ‘adding’ vector spaces together. That is, given two vector spaces V and W we can define their **direct sum** $V \oplus W$ to be the set $V \times W$ with vector addition and scalar multiplication defined by

$$\begin{aligned} (v_1, w_1) + (v_2, w_2) &= (v_1 + v_2, w_1 + w_2) \\ c(v, w) &= (cv, cw). \end{aligned}$$

It is straightforward to check that with vector addition and scalar multiplication so defined, $V \oplus W$ is a bona fide vector space. Also, V can be considered a subset of $V \oplus W$, as just the set of all vectors of the form $(v, 0)$, and likewise for W . With this

identification it is clear that any element $(v, w) \in V \oplus W$ can be written uniquely as $v + w$ with $v \in V$ and $w \in W$, so this notion of a direct sum is consistent with our earlier definition: if we take the direct sum of V and W , the resulting vector space really can be decomposed into the subspaces V and W .

This notion¹³ of direct sum may also be extended to representations; that is, given two representations (Π_i, V_i) , $i = 1, 2$, we may construct the **direct sum representation** $(\Pi_1 \oplus \Pi_2, V_1 \oplus V_2)$ defined by

$$((\Pi_1 \oplus \Pi_2)(g))(v_1, v_2) \equiv (\Pi_1(g)v_1, \Pi_2(g)v_2) \quad \forall (v_1, v_2) \in V_1 \oplus V_2, g \in G.$$

These constructions may seem trivial, but they have immediate physical application, as we will now see.

Example 5.26 The Dirac spinor

Consider the left- and right-handed spinor representations of $SL(2, \mathbb{C})$, (Π, \mathbb{C}^2) and $(\bar{\Pi}, \mathbb{C}^2)$. We define the **Dirac spinor representation** of $SL(2, \mathbb{C})$ to be the direct sum representation $(\Pi \oplus \bar{\Pi}, \mathbb{C}^2 \oplus \mathbb{C}^2)$, which is then given by

$$((\Pi \oplus \bar{\Pi})(A))(v, w) \equiv (Av, A^{\dagger-1}w) \quad \forall (v, w) \in \mathbb{C}^2 \oplus \mathbb{C}^2, A \in SL(2, \mathbb{C}).$$

Making the obvious identification of $\mathbb{C}^2 \oplus \mathbb{C}^2$ with \mathbb{C}^4 , we can write $(\Pi \oplus \bar{\Pi})(A)$ in block matrix form as

$$(\Pi \oplus \bar{\Pi})(A) = \begin{pmatrix} A & 0 \\ 0 & A^{\dagger-1} \end{pmatrix}$$

as in (5.64).

This representation is, by construction, decomposable. Why deal with a decomposable representation rather than its irreducible components? There are a few different ways to answer this in the case of the Dirac spinor, but one rough answer has to do with parity. We will show in Sect. 5.11 that it is impossible to define a consistent action of the parity operator on either the left-handed or right-handed spinors individually, and that what the parity operator naturally wants to do is *interchange* the two representations. Thus, to have a representation of $SL(2, \mathbb{C})$ spinors on which parity naturally acts, we must combine both the left- and right-handed spinors into a Dirac spinor, and in the most natural cases parity is represented by

$$(\Pi \oplus \bar{\Pi})(\text{Parity}) = \pm \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad (5.72)$$

which obviously just interchanges the left- and right-handed spinors. From this we see that the Dirac spinor is reducible under $SL(2, \mathbb{C})$, but not under larger groups which include parity.

¹³In some texts our first notion of direct sum, in which we decompose a vector space into mutually exclusive subspaces, is called an **internal direct sum**, and our second notion of direct sum, in which we take distinct vector spaces and add them together, is known as an **external direct sum**.

5.7 More on Irreducibility

In the last section we introduced the notion of an irreducible representation, but we did not prove that any of the representations we have met are irreducible. In this section we will remedy that and also learn a bit more about irreducibility along the way. In proving the irreducibility of a given representation, the following proposition about the irreps of matrix Lie groups and their Lie algebras is often useful. The proposition just says that for a *connected* matrix Lie group, the irreps of the group are the same as the irreps of the Lie algebra. This may seem unsurprising and perhaps even trivial, but the conclusion does not hold when the group is disconnected. We will have more to say about this later. The proof of this proposition is also a nice exercise in using some of the machinery we have developed so far.

Proposition 5.3 *A representation (Π, V) of a connected matrix Lie group G is irreducible if and only if the induced Lie algebra representation (π, V) of \mathfrak{g} is irreducible as well.*

Proof First, assume that (Π, V) is an irrep of G . Then consider the induced Lie algebra representation (π, V) , and suppose that this representation has an invariant subspace W . We will show that W must be an invariant subspace of (Π, V) as well, which by the irreducibility of (Π, V) will imply that W is either V or $\{0\}$, which will then show that (π, V) is irreducible. For all $X \in \mathfrak{g}$, $w \in W$, we have

$$\begin{aligned} \pi(X)w &\in W \\ \Rightarrow e^{\pi(X)}w &\in W \\ \Rightarrow \Pi(e^X)w &\in W. \end{aligned} \tag{5.73}$$

Now, using the fact that any element of G can be written as a product of exponentials (cf. Proposition 5.2), we then have

$$\begin{aligned} \Pi(g)w &= \Pi(e^{t_1 X_1} e^{t_2 X_2} \dots e^{t_n X_n})w \\ &= \Pi(e^{t_1 X_1})\Pi(e^{t_2 X_2}) \dots \Pi(e^{t_n X_n})w \end{aligned}$$

which must be in W by repeated application of (5.73). Thus W is also an invariant subspace of (Π, V) , but we assumed (Π, V) was irreducible, and so W must be equal to V or 0 . This then proves that (π, V) is irreducible as well.

Conversely, assume that (π, V) is irreducible, and let W be an invariant subspace of (Π, V) . Then for all $t \in \mathbb{R}$, $X \in \mathfrak{g}$, $w \in W$ we know that $\Pi(e^{tX})w \in W$, and hence

$$\begin{aligned} \pi(X)w &= \left. \frac{d}{dt} \Pi(e^{tX})w \right|_{t=0} \\ &= \lim_{t \rightarrow 0} \frac{\Pi(e^{tX})w - w}{t} \end{aligned}$$

must be in W as well, so W is invariant under π and hence must be equal to V or $\{0\}$. Thus (Π, V) is irreducible. \square

Note that the assumption of connectedness was crucial in the above proof, as otherwise we could not have used Proposition 5.2. Furthermore, we will soon meet irreducible representations of disconnected groups (like $O(3, 1)$) that yield Lie algebra representations which *do* have non-trivial invariant subspaces, and are thus not irreducible. The above proposition is still very useful, however, as we will see in this next example.

Example 5.27 *The $SU(2)$ representation on $P_l(\mathbb{C}^2)$, revisited*

In this example we will prove that the $\mathfrak{su}(2)$ representations $(\pi_l, P_l(\mathbb{C}^2))$ of Example 5.7 are all irreducible. Proposition 5.3 will then tell us that the $SU(2)$ representations $(\Pi_l, P_l(\mathbb{C}^2))$ are all irreducible as well. Later on, we will see that these representations are in fact *all* the finite-dimensional irreducible representations of $SU(2)$!

Recall that the $\pi_l(S_i)$ are given by

$$\begin{aligned}\pi_l(S_1) &= \frac{i}{2} \left(z_2 \frac{\partial}{\partial z_1} + z_1 \frac{\partial}{\partial z_2} \right) \\ \pi_l(S_2) &= \frac{1}{2} \left(z_2 \frac{\partial}{\partial z_1} - z_1 \frac{\partial}{\partial z_2} \right) \\ \pi_l(S_3) &= \frac{i}{2} \left(z_1 \frac{\partial}{\partial z_1} - z_2 \frac{\partial}{\partial z_2} \right).\end{aligned}$$

Define now the “raising” and “lowering” operators

$$\begin{aligned}Y_l &\equiv \pi_l(S_2) + i\pi_l(S_1) = -z_1 \frac{\partial}{\partial z_2} \\ X_l &\equiv -\pi_l(S_2) + i\pi_l(S_1) = -z_2 \frac{\partial}{\partial z_1}.\end{aligned}$$

It is easy to see that Y_l trades a factor of z_2 for a factor of z_1 , thereby raising the $\pi_l(S_3)$ eigenvalue of a single term by i . Likewise, X_l trades a z_1 for a z_2 and lowers the eigenvalue by i .¹⁴ Consider a nonzero invariant subspace $W \subset P_l(\mathbb{C}^2)$. If we can show that $W = P_l(\mathbb{C}^2)$, then we know that $P_l(\mathbb{C}^2)$ is irreducible. Being nonzero, W contains at least one element of the form

$$w = a_l z_1^l + a_{l-1} z_1^{l-1} z_2 + a_{l-2} z_1^{l-2} z_2^2 + \cdots + a_0 z_2^l$$

where at least one of the a_k is not zero. Let k_0 be the biggest nonzero value of k , so that $a_{k_0} z_1^{k_0} z_2^{l-k_0}$ is the term in w with the highest power of z_1 . Then applying $(X_l)^{k_0}$

¹⁴You may object to the use of i in our definition of these operators; after all, $\mathfrak{su}(2)$ is a real Lie algebra, and so the expression $S_2 + iS_1$ has no meaning as an element of $\mathfrak{su}(2)$, and so one cannot say that, for instance, $Y_l = \pi(S_2 + iS_1)$. Thus X_l and Y_l are not in the image of $\mathfrak{su}(2)$ under π_l . This is a valid objection, and to deal with it one must introduce the notion of the *complexification* of a Lie algebra. A discussion of this here would lead us too far astray from our main goals of applications in physics, however, so we relegate this material to the appendices which you can consult at leisure.

to w lowers the z_1 degree by k_0 , killing all the terms except $a_{k_0} z_1^{k_0} z_2^{l-k_0}$. In fact, one can compute easily that

$$X_l^{k_0} (a_{k_0} z_1^{k_0} z_2^{l-k_0}) = (-1)^{k_0} k_0! a_{k_0} z_2^l.$$

This is proportional to z_2^l , and since W is invariant W must then contain z_2^l . But then we can successively apply the raising operator Y_l to get monomials of the form $z_1^k z_2^{l-k}$ for $0 \leq k \leq l$, and so these must be in W as well. These, however, form a basis for $P_l(\mathbb{C}^2)$, hence W must equal $P_l(\mathbb{C}^2)$, and so $P_l(\mathbb{C}^2)$ is irreducible. \square

Before moving on to our next examples, we need to state and prove Schur's lemma, one of the most basic and crucial facts about irreducible representations. Roughly speaking, the upshot of it is that any linear operator on the representation space of an irrep which commutes with the group action (i.e. is an intertwiner) must be proportional to the identity. The precise statement is as follows:

Proposition 5.4 (Schur's lemma) *Let (Π_i, V_i) , $i = 1, 2$ be two irreducible representations of a group or Lie algebra, and let $\phi : V_1 \rightarrow V_2$ be an intertwiner. Then either $\phi = 0$ or ϕ is a vector space isomorphism. Furthermore, if $(\Pi_1, V_1) = (\Pi_2, V_2)$ and V_1 is a finite-dimensional complex vector space, then ϕ is a multiple of the identity, i.e. $\phi = cI$ for some $c \in \mathbb{C}$.*

Proof We will prove this for group representations Π_i ; the Lie algebra case follows immediately with the obvious notational changes. Let K be the kernel or null space of ϕ . Then $K \subset V_1$ is an invariant subspace of Π_1 , since for any $v \in K$, $g \in G$,

$$\begin{aligned} \phi(\Pi_1(g)v) &= \Pi_2(g)(\phi(v)) \\ &= \Pi_2(0) \\ &= 0 \\ &\Rightarrow \Pi_1(g)v \in K. \end{aligned}$$

However, since (Π_1, V_1) is irreducible, the only invariant subspaces are 0 and V_1 , so K must be one of those. If $K = V_1$ then $\phi = 0$, so henceforth we assume that $K = 0$, which means that ϕ is one-to-one, and so $\phi(V_1) \subset V_2$ is isomorphic to V_1 . Furthermore, $\phi(V_1)$ is an invariant subspace of (Π_2, V_2) , since for any $\phi(v) \in \phi(V_1)$, $g \in G$ we have

$$(\Pi_2(g))(\phi(v)) = \phi(\Pi_1(g)v) \in \phi(V_1).$$

But (Π_2, V_2) is also irreducible, so $\phi(V_1)$ must equal 0 or V_2 . We already assumed that $\phi(V_1) \neq 0$, though, so we conclude that $\phi(V_1) = V_2$ and hence ϕ is an isomorphism.

Now assume that $(\Pi_1, V_1) = (\Pi_2, V_2) \equiv (\Pi, V)$ and that V is a finite-dimensional complex vector space of dimension n (notice that we did not assume finite-dimensionality at the outset of the proof). Then ϕ is a linear operator and the eigenvalue equation

$$\det(\phi - \lambda I)$$

is a n th degree complex polynomial in λ . By the fundamental theorem of algebra,¹⁵ this polynomial has at least one root $c \in \mathbb{C}$, hence $\phi - cI$ has determinant 0 and is thus non-invertible. This means that $\phi - cI$ has a non-trivial kernel K . K is invariant, though; for all $v \in K$,

$$\begin{aligned} (\phi - cI)(\Pi(g)v) &= \phi(\Pi(g)v) - c\Pi(g)v \\ &= \Pi(g)(\phi(v) - cv) \\ &= 0 \\ &\Rightarrow \Pi(g)v \in K \end{aligned}$$

and since we know $K \neq 0$, we then conclude by irreducibility of V that $K = V$. This means that $(\phi - cI)(v) = 0 \forall v \in V$, which means $\phi = cI$, as desired. \square

Exercise 5.33 Prove the following corollary of Schur's lemma: *If (Π_i, V_i) , $i = 1, 2$ are two complex irreducible representations of a group or Lie algebra, and $\phi, \psi : V_1 \rightarrow V_2$ are two intertwiners with $\phi \neq 0$, then $\psi = c\phi$ for some $c \in \mathbb{C}$.*

As a first application of Schur's lemma, we have

Proposition 5.5 *An irreducible finite-dimensional complex representation of an abelian group or Lie algebra is one-dimensional.*

Proof Again we prove only the group case. Since G is abelian, each $\Pi(g)$ commutes with $\Pi(h)$ for all $h \in G$, hence each $\Pi(g) : V \rightarrow V$ is an intertwiner! By Schur's lemma, this implies that every $\Pi(g)$ is proportional to the identity (with possibly varying proportionality constants), and so *every* subspace of V is an invariant one. Thus the only way V could have no non-trivial invariant subspaces is to have no non-trivial subspaces at all, which means it must be one-dimensional. \square

Exercise 5.34 Show that the fundamental representation of $SO(2)$ on \mathbb{R}^2 is irreducible. Prove this by contradiction, showing that if the fundamental representation were reducible then the $SO(2)$ generator

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

would be diagonalizable over the real numbers, which you should show it is not. This shows that one really needs the hypothesis of a *complex* vector space in the above proposition.

Example 5.28 *The irreducible representations of \mathbb{Z}_2*

Proposition 5.5 allows us to easily enumerate all the irreducible representations of \mathbb{Z}_2 . Since \mathbb{Z}_2 is abelian any irreducible representation (Π_{irr}, V) must be one-dimensional (i.e. $V = \mathbb{R}$ or \mathbb{C}), and Π_{irr} must also satisfy

¹⁵See Herstein [9], for instance.

$$(\Pi_{\text{irr}}(-1))^2 = \Pi_{\text{irr}}((-1)^2) = \Pi_{\text{irr}}(1) = 1,$$

which means that $\Pi_{\text{irr}}(-1) = \pm 1$, i.e. Π_{irr} is either the alternating representation or the trivial representation! Furthermore, \mathbb{Z}_2 is semi-simple (as you will show in Problem 5.8), and so *any* representation (Π, V) of \mathbb{Z}_2 is completely reducible and thus decomposes into one-dimensional irreducible subspaces, on which $\Pi(-1)$ equals either 1 or -1 . Thus $\Pi(-1)$ is diagonalizable, with eigenvalues ± 1 (cf. Example 5.18 for examples of this). Let us say the \mathbb{Z}_2 in question is $\mathbb{Z}_2 \simeq \{I, P\} \subset O(3)$, and that we are working in a quantum mechanical context with a Hilbert space \mathcal{H} . Then there exists a basis for \mathcal{H} consisting of eigenvectors of $\Pi(P)$; for a given eigenvector ψ , its eigenvalue of ± 1 is known as its *parity*. If the eigenvalue is $+1$ then ψ is said to have *even* parity, and if the eigenvalue is -1 then ψ is said to have *odd* parity. \square

5.8 The Irreducible Representations of $\mathfrak{su}(2)$, $SU(2)$ and $SO(3)$

In this section we will construct (up to equivalence) *all* the finite-dimensional irreducible complex representations of $\mathfrak{su}(2)$. Besides being of intrinsic interest, our results will also allow us to classify all the irreducible representations of $SO(3)$, $SU(2)$, and even the apparently unrelated representations of $\mathfrak{so}(3, 1)$, $SO(3, 1)$ and $SL(2, \mathbb{C})$. The construction we will give is more or less the same as that found in the physics literature under the heading “theory of angular momentum”, except that we are using different language and notation. Our strategy will be to use the commutation relations to deduce the possible structures of $\mathfrak{su}(2)$ irreps, and then show that we have already constructed representations which exhaust these possibilities, thus yielding a complete classification.

Let (π, V) be a finite-dimensional complex irreducible representation of $\mathfrak{su}(2)$. It will be convenient to use the following shorthand, familiar from the physics literature:

$$\begin{aligned} J_z &\equiv i\pi(S_z) \\ J_+ &\equiv i\pi(S_x) - \pi(S_y) \\ J_- &\equiv i\pi(S_x) + \pi(S_y). \end{aligned} \tag{5.74}$$

These “raising” and “lowering” operators obey the following commutation relations, as you can check:

$$\begin{aligned} [J_z, J_\pm] &= \pm J_\pm \\ [J_+, J_-] &= 2J_z. \end{aligned}$$

Now, as discussed in our proof of Schur’s lemma, the fact that V is complex means that every operator on V has at least one eigenvector. In particular, this means that J_z has an eigenvector v with eigenvalue b . The above commutation relations then imply that

$$J_z(J_{\pm}v) = [J_z, J_{\pm}]v + J_{\pm}(J_zv) = (b \pm 1)J_{\pm}v$$

so that if $J_{\pm}v$ is not zero (which it might be!), then it is another eigenvector of J_z with eigenvalue $b \pm 1$. Now, we can repeatedly apply J_+ to v to yield more and more eigenvectors of J_z , but since V is finite-dimensional and eigenvectors with different eigenvalues are linearly independent (see Exercise 5.35 below), this process must end somewhere, say at N applications of J_+ . Let v_0 be this vector with the highest eigenvalue (also known as the **highest weight vector** of the representation), so that we have

$$\begin{aligned} v_0 &= (J_+)^N v \\ J_+ v_0 &= 0. \end{aligned}$$

Then v_0 has a J_z eigenvalue of $b + N \equiv j$ (note that so far we haven't proved anything about b or j , but we will soon see that they must be integral or half-integral). Starting with v_0 , then, we can repeatedly 'lower' with J_- to get eigenvectors with lower eigenvalues. In fact, we can define

$$v_k \equiv (J_-)^k v_0$$

which has J_z eigenvalue $j - k$. This chain must also end, though, so there must exist an integer l such that

$$v_l = (J_-)^l v_0 \neq 0 \quad \text{but} \quad v_{l+1} = (J_-)^{l+1} v_0 = 0.$$

How can we find l ? For this we will need the following formula, which you will prove in Exercise 5.36 below:

$$J_+(v_k) = [2jk - k(k-1)]v_{k-1}. \quad (5.75)$$

Applying this to $v_{l+1} = 0$ gives

$$0 = J_+(v_{l+1}) = [2j(l+1) - (l+1)l]v_l$$

and since $v_l \neq 0$ we conclude that

$$[2j(l+1) - (l+1)l] = 0 \quad \Longleftrightarrow \quad j = l/2.$$

Thus j is a nonnegative integer or half integer! (In fact, j is just the 'spin' of the representation.) We further conclude that V contains $2j + 1$ vectors $\{v_k \mid 0 \leq k \leq 2j\}$ all of which are eigenvectors of J_z with eigenvalue $j - k$. Furthermore, since

$$\begin{aligned} \pi(S_x) &= -\frac{i}{2}(J_+ + J_-) \\ \pi(S_y) &= \frac{1}{2}(J_- - J_+), \end{aligned}$$

the action of S_x and S_y take a given v_k into a linear combination of other v_k , so the span of v_k is a nonzero invariant subspace of V . We assumed V was irreducible, however, so we must have $V = \text{Span}\{v_k\}$, and since the v_k are linearly independent, they form a basis for V !

To summarize, any finite-dimensional irreducible complex representation of $\mathfrak{su}(2)$ has dimension $2j + 1$, $2j \in \mathbb{Z}$ and a basis $\{v_k\}_{k=0-2j}$ which satisfies

$$\begin{aligned} J_+(v_0) &= 0. \\ J_-(v_k) &= v_{k+1}, \quad k < 2j \\ J_z(v_k) &= (j - k)v_k \\ J_-(v_{2j}) &= 0 \\ J_+(v_k) &= [2jk - k(k - 1)]v_{k-1}, \quad k \neq 0. \end{aligned} \tag{5.76}$$

What is more, we can actually use the above equations to *define* representations (π_j, V_j) , where V_j is a $(2j + 1)$ -dimensional vector space with basis $\{v_k\}_{k=0-2j}$ and the action of the operators $\pi_j(S_i)$ is defined by (5.76). It is straightforward to check that this defines a representation of $\mathfrak{su}(2)$ (see exercise below), and one can prove irreducibility in the same way that we did in Example 5.27. Furthermore, any irrep (π, V) of $\mathfrak{su}(2)$ must be equivalent to (π_j, V_j) for some j , since we can find a basis w_k for V satisfying (5.76) for some j and then define an intertwiner by

$$\begin{aligned} \phi : V &\rightarrow V_j \\ w_k &\mapsto v_k \end{aligned}$$

and extending linearly. We have thus proved the following:

Proposition 5.6 *The $\mathfrak{su}(2)$ representations (π_j, V_j) , $2j \in \mathbb{Z}$ defined above are all irreducible, and any other finite-dimensional complex irreducible representation of $\mathfrak{su}(2)$ is equivalent to (π_j, V_j) for some j , $2j \in \mathbb{Z}$.*

In other words, the (π_j, V_j) are, up to equivalence, *all* the finite-dimensional complex irreducible representations of $\mathfrak{su}(2)$. They are also all the finite-dimensional complex irreducible representations of $\mathfrak{so}(3)$, since $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$.

If you look back over our arguments you will see that we deduced (5.76) from just the $\mathfrak{su}(2)$ commutation relations, the finite-dimensionality of V , and the existence of a highest weight vector v_0 satisfying $J_+(v_0) = 0$ and $J_z(v_0) = jv_0$. Thus, if we have an arbitrary (i.e. not necessarily irreducible) finite-dimensional $\mathfrak{su}(2)$ representation (π, V) and can find a highest weight vector v_0 for some j , we can lower with J_- to generate a basis $\{v_k\}_{k=0-2j}$ satisfying (5.76) and conclude that V has an invariant subspace equivalent to (π_j, V_j) . We can then repeat this until V is completely decomposed into irreps. If we know that (π, V) is irreducible from the start then we do not even have to find the vector v_0 , we just use the fact that (π, V) must be equivalent to (π_j, V_j) for some j and note that j is given by $j = \frac{1}{2}(\dim V - 1)$. These observations make it easy to identify which (π_j, V_j) occur in any given $\mathfrak{su}(2)$ representation.

Note that if we have a finite-dimensional complex irreducible $\mathfrak{su}(2)$ representation that is also *unitary*, then we could work in an orthonormal basis. In that case, it turns out that the v_k defined above are *not* orthonormal, and are thus not ideal basis vectors to work with. They are orthogonal, but are not normalized to have

unit length. In quantum mechanical contexts the $\mathfrak{su}(2)$ representations usually are unitary, and so in that setting one works with the orthonormal basis vectors $|m\rangle$, $-j \leq m \leq j$. The vector $|m\rangle$ is proportional to our v_{j-m} , and is normalized whereas v_{j-m} is not. See Sakurai [14] for details on the normalization procedure.

Exercise 5.35 Let $S = \{v_i\}_{i=1-k}$ be a set of eigenvectors of some linear operator T on a vector space V . Show that if each of the v_i have *distinct* eigenvalues, then S is a linearly independent set. (Hint: One way to do this is by induction on k . Another is to argue by assuming that S is linearly dependent and reaching a contradiction. In this case you may assume without loss of generality that the v_i , $1 \leq i \leq k-1$ are linearly independent, so that v_k is the vector that spoils the assumed linear independence.)

Exercise 5.36 Prove (5.75). Proceed by induction, i.e. first prove the formula for $k = 1$, then assume it is true for k and show that it must be true for $k + 1$.

Exercise 5.37 Show that (5.76) defines a representation of $\mathfrak{su}(2)$. This consists of showing that the operators J_+ , J_- , J_z satisfy the appropriate commutation relations. Then show that this representation is irreducible, using an argument similar to the one from Example 5.27.

Example 5.29 $P_l(\mathbb{C}^2)$, revisited again

In Example 5.7 we met the representations $(\pi_l, P_l(\mathbb{C}^2))$ of $\mathfrak{su}(2)$ on the space of degree l polynomials in two complex variables. In Example 5.27 we saw that these representations are all irreducible, and so by setting $\dim P_l(\mathbb{C}^2) = l + 1$ equal to $2j + 1$ we deduce that

$$\boxed{(\pi_l, P_l(\mathbb{C}^2)) \simeq (\pi_{l/2}, V_{l/2})} \quad (5.77)$$

and so the $(\pi_l, P_l(\mathbb{C}^2))$, $l \in \mathbb{Z}$ also yield all the complex finite-dimensional irreps of $\mathfrak{su}(2)$. What is more, this allows us to enumerate all the finite-dimensional complex irreps of the associated group $SU(2)$. Any irrep (Π, V) of $SU(2)$ yields an irrep (π, V) of $\mathfrak{su}(2)$, by Proposition 5.3. This irrep must be equivalent to $(\pi_l, P_l(\mathbb{C}^2))$ for some $l \in \mathbb{Z}$, however, and so by Proposition 5.2 (Π, V) is equivalent to $(\Pi_l, P_l(\mathbb{C}^2))$. Thus, **the representations $(\Pi_l, P_l(\mathbb{C}^2))$, $l \in \mathbb{Z}$ are (up to equivalence) all the finite-dimensional complex irreducible representations of $SU(2)$!**

It is instructive to construct the equivalence (5.77) explicitly. Recall that the raising and lowering operators (which we called Y_l and X_l in Example 5.27) are given by

$$J_- = -z_1 \frac{\partial}{\partial z_2}$$

$$J_+ = -z_2 \frac{\partial}{\partial z_1}$$

and that

$$J_z = i\pi_l(S_z) = \frac{1}{2} \left(z_2 \frac{\partial}{\partial z_2} - z_1 \frac{\partial}{\partial z_1} \right).$$

It is easy to check that $v_0 \equiv z_2^l$ is a highest weight vector with $j = l/2$, and so the basis that satisfies (5.76) is given by

$$v_k \equiv (J_-)^k (z_2^l) = (-1)^k \frac{l!}{(l-k)!} z_1^k z_2^{l-k}, \quad 0 \leq k \leq l. \quad (5.78)$$

Exercise 5.38 Show by direct calculation that $(\pi_{2j}, P_{2j}(\mathbb{C}^2))$ satisfies (5.76) with basis vectors given by (5.78).

Example 5.30 $S^{2j}(\mathbb{C}^2)$ as irreps of $\mathfrak{su}(2)$

We suggested in Example 5.15 that the tensor product representation of $\mathfrak{su}(2)$ on $S^{2j}(\mathbb{C}^2)$, the totally symmetric $(0, 2j)$ tensors on \mathbb{C}^2 , is equivalent to $(\pi_{2j}, P_{2j}(\mathbb{C}^2)) \simeq (\pi_j, V_j)$, $2j \in \mathbb{Z}$. With the classification of $\mathfrak{su}(2)$ irreps in place, we can now prove this fact. It is easy to verify that

$$v_0 \equiv \underbrace{e_1 \otimes \cdots \otimes e_1}_{2j \text{ times}} \in S^{2j}(\mathbb{C}^2) \quad (5.79)$$

is a highest weight vector with eigenvalue j , so there is an irreducible invariant subspace of $S^{2j}(\mathbb{C}^2)$ equivalent to V_j . Since $\dim V_j = \dim S^{2j}(\mathbb{C}^2) = 2j + 1$ (as you can check), we conclude that V_j is all of $S^{2j}(\mathbb{C}^2)$, and hence that

$$\boxed{(S^{2j}\pi, S^{2j}(\mathbb{C}^2)) \simeq (\pi_j, V_j) \simeq (\pi_{2j}, P_{2j}(\mathbb{C}^2))} \quad (5.80)$$

which also implies $(S^{2j}\pi, S^{2j}(\mathbb{C}^2)) \simeq (\pi_{2j}, P_{2j}(\mathbb{C}^2))$. Thus, we see that **every irreducible representation of $\mathfrak{su}(2)$ and $SU(2)$ can be obtained by taking a symmetric tensor product of the fundamental representation**. Thus, using nothing more than the fundamental representation (which corresponds to $j = 1/2$, as expected) and the tensor product, we can generate all the representations of $\mathfrak{su}(2)$ and $SU(2)$. We will see in the next section that the same is true for $\mathfrak{so}(3, 1)$ and $SO(3, 1)_0$. \square

Our results about the finite-dimensional irreps of $\mathfrak{su}(2)$ and $SU(2)$ are not only interesting in their own right; they also allow us to determine all the irreps of $SO(3)$! To see this, first consider the degree l harmonic polynomial representation $(\pi_l, \mathcal{H}_l(\mathbb{R}^3))$ of $SO(3)$. This induces a representation $(\pi_l, \mathcal{H}_l(\mathbb{R}^3))$ of $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$. It is easy to check (see exercise below) that if we define, in analogy to the $\mathfrak{su}(2)$ case,

$$\begin{aligned} J_z &\equiv i\pi_l(L_z) = i\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) \\ J_+ &\equiv i\pi(L_x) - \pi(L_y) = i\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) - \left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right) \\ J_- &\equiv i\pi(L_x) + \pi(L_y) = i\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right) + \left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right) \end{aligned}$$

then the vector $f_0 \equiv (x + iy)^l$ is a highest weight vector with eigenvalue l , and so $\mathcal{H}_l(\mathbb{R}^3)$ has an invariant subspace equivalent to (π_l, V_l) . We will argue in Problem 5.9 that $\dim \mathcal{H}_l(\mathbb{R}^3) = 2l + 1$, so we conclude that $(\pi_l, \mathcal{H}_l(\mathbb{R}^3)) \simeq (\pi_l, V_l)$, and

is hence an irrep of $\mathfrak{so}(3)$. Proposition 5.3 then implies that $(\Pi_l, \mathcal{H}_l(\mathbb{R}^3))$ is an irrep of $SO(3)$!

Are these *all* the irreps of $SO(3)$? To find out, let (Π, V) be an arbitrary finite-dimensional complex irrep of $SO(3)$. Then the induced $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$ representation (π, V) must be equivalent to (π_j, V_j) for some integral or half-integral j . We just saw that any integral j value is possible, by taking $(\Pi, V) = (\Pi_j, \mathcal{H}_j(\mathbb{R}^3))$. What about half-integral values of j ? In this case, we have (careful not to confuse the number π with the representation π !)

$$\begin{aligned} e^{2\pi \cdot \pi(L_z)} v_0 &= e^{-i2\pi J_z} v_0 \\ &= e^{-i2\pi j} v_0 \quad \text{since } v_0 \text{ has eigenvalue } j \\ &= -v_0 \quad \text{since } j \text{ half-integral.} \end{aligned}$$

However, we also have

$$\begin{aligned} e^{2\pi \cdot \pi(L_z)} v_0 &= (\Pi(e^{2\pi \cdot L_z})) v_0 \\ &= (\Pi(I)) v_0 \quad \text{by (4.86)} \\ &= v_0, \end{aligned}$$

a contradiction. Thus, j cannot be half-integral, so (π, V) must be equivalent to (π_j, V_j) for j integral. But this implies that $(\pi, V) \simeq (\pi_j, \mathcal{H}_j(\mathbb{R}^3))$, and so by Proposition 5.2, (Π, V) is equivalent to $(\Pi_j, \mathcal{H}_j(\mathbb{R}^3))$! Thus, **the representations $(\Pi_j, \mathcal{H}_j(\mathbb{R}^3))$, $j \in \mathbb{Z}$ are (up to equivalence) all the finite-dimensional complex irreducible representations of $SO(3)$.**

An important lesson to take away from this is that for a matrix Lie group G with Lie algebra \mathfrak{g} , **not all representations of \mathfrak{g} necessarily come from representations of G .** If $G = SU(2)$ then there is a one-to-one correspondence between Lie algebra representations and group representations, but in the case of $SO(3)$ there are Lie algebra representations (corresponding to half-integral values of j) that *do not* come from $SO(3)$ representations. We will not say much more about this here, except to note that this is connected to the fact that there are non-isomorphic matrix Lie groups that have isomorphic Lie algebras, as is the case with $SU(2)$ and $SO(3)$. For a more complete discussion, see Hall [8].

Exercise 5.39 Verify that

$$\begin{aligned} J_+ f_0 &= 0 \\ J_z f_0 &= l f_0. \end{aligned}$$

5.9 Real Representations and Complexifications

So far we have classified all the *complex* finite-dimensional irreps of $\mathfrak{su}(2)$, $SU(2)$ and $SO(3)$, but we haven't said anything about *real* representations, despite the fact that many of the most basic representations of these groups and Lie algebras (like the fundamental of $SO(3)$ and all its various tensor products) are real. Fortunately, there is a way to turn every real representation into a complex representation, so

that we can then apply our classification of complex irreps. Given any real vector space V , we can define the **complexification** of V as $V_{\mathbb{C}} \equiv \mathbb{C} \otimes V$, where \mathbb{C} and V are thought of as real vector spaces (\mathbb{C} being a 2-dimensional real vector space with basis $\{1, i\}$), so that if $\{e_i\}$ is a basis for V then $\{1 \otimes e_i, i \otimes e_i\}$ is a (real) basis for $V_{\mathbb{C}}$. Note that $V_{\mathbb{C}}$ also carries the structure of a *complex* vector space, with multiplication by i defined by

$$i(z \otimes v) = (iz) \otimes v, \quad z \in \mathbb{C}, v \in V.$$

A complex basis for $V_{\mathbb{C}}$ is then given by $\{1 \otimes e_i\}$, and the complex dimension of $V_{\mathbb{C}}$ is equal to the real dimension of V .

We can then define the **complexification** of a real representation (Π, V) to be the (complex) representation $(\Pi_{\mathbb{C}}, V_{\mathbb{C}})$ defined by

$$(\Pi_{\mathbb{C}}(g))(z \otimes v) \equiv z \otimes \Pi(g)v. \quad (5.81)$$

We can then get a handle on the real representation (Π, V) by applying our classification scheme to its complexification $(\Pi_{\mathbb{C}}, V_{\mathbb{C}})$. You should be aware, however, that the irreducibility of (Π, V) does not guarantee the irreducibility of $(\Pi_{\mathbb{C}}, V_{\mathbb{C}})$ (see Exercise 5.40 and Example 5.39), though in many cases $(\Pi_{\mathbb{C}}, V_{\mathbb{C}})$ will end up being irreducible.

Example 5.31 *The complexifications of \mathbb{R}^n and $M_n(\mathbb{R})$*

As a warm-up to considering complexifications of representations, we consider the complexification of a simple vector space. Consider the complexification $\mathbb{R}_{\mathbb{C}}^n$ of \mathbb{R}^n . We can define the (obvious) complex-linear map

$$\begin{aligned} \phi : \mathbb{R}_{\mathbb{C}}^n &\rightarrow \mathbb{C}^n \\ 1 \otimes (a^j e_j) + i \otimes (b^j e_j) &\mapsto (a^j + ib^j)e_j, \quad a^j, b^j \in \mathbb{R} \end{aligned}$$

which is easily seen to be a vector space isomorphism, so we can identify $\mathbb{R}_{\mathbb{C}}^n$ with \mathbb{C}^n . One can also extend this argument in the obvious way to show that

$$(M_n(\mathbb{R}))_{\mathbb{C}} = M_n(\mathbb{C}).$$

Example 5.32 *The fundamental representation of $\mathfrak{so}(3)$*

Now consider the complexification $(\pi_{\mathbb{C}}, \mathbb{R}_{\mathbb{C}}^3)$ of the fundamental representation of $\mathfrak{so}(3)$. As explained above, $\mathbb{R}_{\mathbb{C}}^3$ can be identified with \mathbb{C}^3 , and a moment's consideration of (5.81) will show that the complexification of the fundamental representation of $\mathfrak{so}(3)$ is just given by the usual $\mathfrak{so}(3)$ matrices acting on \mathbb{C}^3 rather than \mathbb{R}^3 . You should check that J_z and J_+ are given by

$$J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$J_+ = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -i \\ 1 & i & 0 \end{pmatrix}$$

and that the vector

$$v_0 \equiv e_1 + ie_2 = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$$

is a highest weight vector with $j = 1$. This, along with the fact that $\dim \mathbb{C}^3 = 3$, allows us to conclude that $(\pi_{\mathbb{C}}, \mathbb{R}_{\mathbb{C}}^3) \simeq (\pi_1, V_1)$. This is why regular three-dimensional vectors are said to be “spin-one”. Notice that our highest weight vector $v_0 = e_1 + ie_2$ is just the analog of the function $f_0 = x + iy \in \mathcal{H}_1(\mathbb{R}^3)$, which is an equivalent representation, and that in terms of $SO(3)$ reps we have

$$(\Pi_{\mathbb{C}}, \mathbb{R}_{\mathbb{C}}^3) \simeq (\Pi_1, \mathcal{H}_1(\mathbb{R}^3)).$$

Recall also that the adjoint representation of $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$ is equivalent to the fundamental representation of $\mathfrak{so}(3)$. This combined with the above results implies

$$(\mathrm{ad}_{\mathbb{C}}, \mathfrak{so}(3)_{\mathbb{C}}) \simeq (\mathrm{ad}_{\mathbb{C}}, \mathfrak{su}(2)_{\mathbb{C}}) \simeq (\pi_1, V_1)$$

so the adjoint representation of $\mathfrak{su}(2)$ is “spin-one” as well.

Example 5.33 *Symmetric traceless tensors*

Consider the space $S'^2(\mathbb{R}^3)$ of symmetric traceless second rank tensors defined in (5.67). This is an $SO(3)$ invariant subspace of $\mathbb{R}^3 \otimes \mathbb{R}^3$, and so furnishes a representation $(\Pi_2^0, S'^2(\mathbb{R}^3))$ of $SO(3)$. What representation is this? To find out, consider the complexification of the associated $\mathfrak{so}(3)$ representation, $(\pi_{0\mathbb{C}}^{2'}, S'^2(\mathbb{C}^3))$, where $S'^2(\mathbb{C}^3)$ is also defined by (5.67), just with \mathbb{C}^3 replacing \mathbb{R}^3 . Then it is straightforward to verify that

$$v_0 \equiv (e_1 + ie_2) \otimes (e_1 + ie_2)$$

is a highest weight vector with $j = 2$, and this along with fact that $\dim S'^2(\mathbb{C}^3) = 5$ (check!) implies that

$$(\pi_{0\mathbb{C}}^{2'}, S'^2(\mathbb{C}^3)) \simeq (\pi_2, V_2).$$

This is why symmetric traceless second rank tensors on \mathbb{R}^3 are sometimes said to be “spin-two”.

Recall that $S'^2(\mathbb{R}^3)$ arose in Example 5.23 as part of the decomposition

$$\mathbb{R}^3 \otimes \mathbb{R}^3 = \mathbb{R}g \oplus \Lambda^2(\mathbb{R}^3) \oplus S'^2(\mathbb{R}^3)$$

which has matrix counterpart

$$M_3(\mathbb{R}) = \mathbb{R}I \oplus A_3(\mathbb{R}) \oplus S'_3(\mathbb{R}).$$

Complexifying and using the fact (cf. Example 5.22) that $\Lambda^2(\mathbb{R}^3)$ is equivalent to the adjoint representation and is hence “spin-one”, we obtain the following decomposition:

$$V_1 \otimes V_1 \simeq \mathbb{C}^3 \otimes \mathbb{C}^3 \simeq M_3(\mathbb{C}) \simeq V_0 \oplus V_1 \oplus V_2.$$

This is an instance of (5.68) and should be familiar from angular momentum addition in quantum mechanics.

Exercise 5.40 Consider the fundamental representation of $SO(2)$ on \mathbb{R}^2 , which we know is irreducible by Exercise 5.34. The complexification of this representation is just given by the same $SO(2)$ matrices acting on \mathbb{C}^2 rather than \mathbb{R}^2 . Show that this representation is reducible, by diagonalizing the $SO(2)$ generator

$$X = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Note that this diagonalization can now be done because both complex eigenvalues and complex basis transformations are allowed, in contrast to the real case.

5.10 The Irreducible Representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, $SL(2, \mathbb{C})$ and $SO(3, 1)_o$

In this section we will use the techniques and results of the previous section to classify all the finite-dimensional complex irreps of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}} \simeq \mathfrak{so}(3, 1)$, and then use this to find the irreps of the associated groups $SL(2, \mathbb{C})$ and $SO(3, 1)_o$.

Let (π, V) be a finite-dimensional complex irreducible representation of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$. Define the operators

$$\boxed{\begin{aligned} M_i &\equiv \frac{1}{2}(\pi(S_i) - i\pi(\tilde{K}_i)), & i = 1, 2, 3 \\ N_i &\equiv \frac{1}{2}(\pi(S_i) + i\pi(\tilde{K}_i)), & i = 1, 2, 3 \end{aligned}} \quad (5.82)$$

where $\{S_i, \tilde{K}_i\}_{i=1,2,3}$ is our usual basis for $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$. One can check that the M s and N s commute between each other, as well as satisfy the $\mathfrak{su}(2)$ commutation relations internally, i.e.

$$\begin{aligned} [M_i, N_j] &= 0 \\ [M_i, M_j] &= \sum_{k=1}^3 \epsilon_{ijk} M_k \\ [N_i, N_j] &= \sum_{k=1}^3 \epsilon_{ijk} N_k. \end{aligned} \quad (5.83)$$

We have thus taken the *complex* span of the set $\{\pi(\tilde{K}_i), \pi(S_i)\}$ (notice the factors of i in (5.82)) and found a new basis for this Lie algebra of operators that makes it

look like two *commuting* copies of $\mathfrak{su}(2)$. We can thus define the usual raising and lowering operators

$$\begin{aligned} N_{\pm} &\equiv iN_1 \mp N_2 \\ M_{\pm} &\equiv iM_1 \mp M_2 \end{aligned}$$

which then have the usual commutation relations between themselves and iM_z, iN_z :

$$\begin{aligned} [iM_z, M_{\pm}] &= \pm M_{\pm} \\ [iN_z, N_{\pm}] &= \pm N_{\pm} \\ [M_+, M_-] &= 2iM_z \\ [N_+, N_-] &= 2iN_z. \end{aligned}$$

With this machinery set up we can now use the strategy from the last section. First, pick a vector $v \in V$ that is an eigenvector of both iM_z and iN_z (that such a vector exists is guaranteed by Problem 5.10). Then by applying M_+ and N_+ we can raise the iM_z and iN_z eigenvalues until the raising operators give us the zero vector; let $v_{0,0}$ denote the vector with the highest eigenvalues (which we will again refer to as a ‘highest weight vector’), and let (j_1, j_2) denote those eigenvalues under iM_z and iN_z , respectively, so that

$$\begin{aligned} M_+(v_{0,0}) &= 0 \\ N_+(v_{0,0}) &= 0 \\ iM_z(v_{0,0}) &= j_1 v_{0,0} \\ iN_z(v_{0,0}) &= j_2 v_{0,0}. \end{aligned} \tag{5.84}$$

We can then lower the eigenvalues with N_- and M_- to get vectors

$$v_{k_1, k_2} \equiv (M_-)^{k_1} (N_-)^{k_2} v_{0,0} \tag{5.85}$$

which are eigenvectors of iM_z and iN_z with eigenvalues $j_1 - k_1$ and $j_2 - k_2$, respectively. By finite-dimensionality of V this chain of vectors must eventually end, though, so there exist nonnegative integers l_1, l_2 such that $v_{l_1, l_2} \neq 0$ but

$$M_-(v_{l_1, l_2}) = N_-(v_{l_1, l_2}) = 0. \tag{5.86}$$

Calculations identical to those from the $\mathfrak{su}(2)$ case show that $l_i = 2j_i$, and that the action of the operators M_i, N_i is given by

$$\begin{aligned} M_+(v_{0,0}) &= 0 \\ N_+(v_{0,0}) &= 0 \\ M_-(v_{k_1, k_2}) &= v_{k_1+1, k_2}, \quad k_1 < 2j_1 \\ N_-(v_{k_1, k_2}) &= v_{k_1, k_2+1}, \quad k_2 < 2j_2 \\ iM_z(v_{k_1, k_2}) &= (j_1 - k_1)v_{k_1, k_2} \\ iN_z(v_{k_1, k_2}) &= (j_2 - k_2)v_{k_1, k_2} \\ M_-(v_{2j_1, k_2}) &= 0, \quad \forall k_2 \end{aligned} \tag{5.87}$$

$$\begin{aligned}
N_-(v_{k_1, 2j_2}) &= 0, \quad \forall k_1 \\
M_+(v_{k_1, k_2}) &= [2j_1 k_1 - k_1(k_1 - 1)]v_{k_1-1, k_2}, \quad k_1 \neq 0 \\
N_+(v_{k_1, k_2}) &= [2j_2 k_2 - k_2(k_2 - 1)]v_{k_1, k_2-1}, \quad k_2 \neq 0
\end{aligned}$$

(this is just two copies of (5.76), one each for the M_i and the N_i). As in the $\mathfrak{su}(2)$ case, we note that the $\{v_{k_1, k_2}\}$ are linearly independent and span an invariant subspace of V , hence must span all of V since we assumed V was irreducible. Thus, we conclude that any complex finite-dimensional irrep of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is of the form $(\pi_{(j_1, j_2)}, V_{(j_1, j_2)})$ where $V_{(j_1, j_2)}$ has a basis

$$\mathcal{B} = \{v_{k_1, k_2} \mid 0 \leq k_1 \leq 2j_1, 0 \leq k_2 \leq 2j_2\}$$

and the operators $\pi_{(j_1, j_2)}(S_i)$, $\pi_{(j_1, j_2)}(\tilde{K}_i)$ satisfy (5.87), with $2j_1, 2j_2 \in \mathbb{Z}$. This tells us that

$$\dim V_{(j_1, j_2)} = (2j_1 + 1)(2j_2 + 1).$$

Let us abbreviate the representations $(\pi_{(j_1, j_2)}, V_{(j_1, j_2)})$ as simply $(\mathbf{j}_1, \mathbf{j}_2)$ as is done in the physics literature. As in the $\mathfrak{su}(2)$ case, we can show that (5.87) actually defines a representation of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, and using the same arguments that we did in the $\mathfrak{su}(2)$ case we conclude that

Proposition 5.7 *The representations $(\mathbf{j}_1, \mathbf{j}_2)$, $2j_1, 2j_2 \in \mathbb{Z}$ are, up to equivalence, all the complex finite-dimensional irreducible representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$.*

As in the $\mathfrak{su}(2)$ case, we deduced (5.87) from just the $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ commutation relations, the finite-dimensionality of V , and the existence of a highest weight vector $v_{0,0}$ satisfying (5.84). Thus if we are given a finite-dimensional $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representation and can find a highest weight vector $v_{0,0}$ for some (j_1, j_2) , we can conclude that the representation space contains an invariant subspace equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$.

Example 5.34 (π, \mathbb{C}^2) : *The fundamental (left-handed spinor) representation*

Consider the left-handed spinor representation (π, \mathbb{C}^2) of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, which is also just the fundamental of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, i.e.

$$\begin{aligned}
\pi(S_1) = S_1 &= \frac{1}{2} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\
\pi(S_2) = S_2 &= \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
\pi(S_3) = S_3 &= \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\
\pi(\tilde{K}_1) = \tilde{K}_1 &= \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\pi(\tilde{K}_2) = \tilde{K}_2 &= \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}
\end{aligned}$$

$$\pi(\tilde{K}_3) = \tilde{K}_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In this case we have (check!)

$$M_i = S_i$$

$$N_i = 0$$

and hence

$$\begin{aligned} M_+ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ M_- &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ N_+ &= N_- = 0. \end{aligned} \tag{5.88}$$

With this in hand you can easily check that $(1, 0) \in \mathbb{C}^2$ is a highest weight vector with $j_1 = 1/2$, $j_2 = 0$. Thus the fundamental representation of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is just $(\frac{1}{2}, \mathbf{0})$.

Example 5.35 ($S^{2j}\pi, S^{2j}(\mathbb{C}^2)$): *Symmetric tensor products of left-handed spinors*

As with $\mathfrak{su}(2)$, we can build other irreps by taking symmetric tensor products. Consider the $2j$ th symmetric tensor product representation $(S^{2j}\pi, S^{2j}(\mathbb{C}^2))$, $2j \in \mathbb{Z}$ and the vector

$$v_{0,0} \equiv \underbrace{e_1 \otimes \cdots \otimes e_1}_{2j \text{ times}} \in S^{2j}(\mathbb{C}^2). \tag{5.89}$$

Using (5.34) it is straightforward to check, as you did in Example 5.30, that this is a highest weight vector with eigenvalues $(j, 0)$, and so we conclude that $S^{2j}(\mathbb{C}^2)$ contains an invariant subspace equivalent to $(\mathbf{j}, \mathbf{0})$. Noting that

$$\dim S^{2j}(\mathbb{C}^2) = \dim(\mathbf{j}, \mathbf{0}) = 2j + 1$$

we conclude that $(S^{2j}\pi, S^{2j}(\mathbb{C}^2)) \simeq (\mathbf{j}, \mathbf{0})$.

Example 5.36 $(\bar{\pi}, \mathbb{C}^2)$: *The right-handed spinor representation*

Consider the right-handed spinor representation $(\bar{\pi}, \mathbb{C}^2)$ from Example 5.6. A quick calculation (do it!) reveals that the induced Lie algebra representation $(\bar{\pi}, \mathbb{C}^2)$ is given by

$$\bar{\pi}(X) = -X^\dagger, \quad X \in \mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}.$$

In particular we then have

$$\bar{\pi}(S_i) = S_i$$

since the S_i are anti-Hermitian, as well as

$$\bar{\pi}(\tilde{K}_i) = \bar{\pi}(iS_i) = -i\bar{\pi}(S_i) = -iS_i.$$

We then have

$$\begin{aligned} M_i &= 0 \\ N_i &= S_i \end{aligned}$$

and hence

$$\begin{aligned} N_+ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ N_- &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \\ M_+ &= M_- = 0. \end{aligned} \tag{5.90}$$

You can again check that $(1, 0) \in \mathbb{C}^2$ is a highest weight vector, but this time with $j_1 = 0$, $j_2 = 1/2$, and so the right-handed spinor representation of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ is just $(\mathbf{0}, \frac{1}{2})$.

Example 5.37 $(\bar{\pi}_{2j}, S^{2j}(\mathbb{C}^2))$: *Symmetric tensor products of right-handed spinors*

As before, we can build other irreps by taking symmetric tensor products. Again, consider the $2j$ th symmetric tensor product representation $(S^{2j}\bar{\pi}, S^{2j}(\mathbb{C}^2))$, $2j \in \mathbb{Z}$ and the vector

$$v_{0,0} \equiv \underbrace{e_1 \otimes \cdots \otimes e_1}_{2j \text{ times}} \in S^{2j}(\mathbb{C}^2). \tag{5.91}$$

Again, it is straightforward to check that this is a highest weight vector with eigenvalues $(0, j)$, and we can conclude as before that $(S^{2j}\bar{\pi}, S^{2j}(\mathbb{C}^2))$ is equivalent to $(\mathbf{0}, \mathbf{j})$. \square

So far we have used symmetric tensor products of the left-handed and right-handed spinor representations to build the $(\mathbf{j}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{j})$ irreps. From here, getting the general irrep $(\mathbf{j}_1, \mathbf{j}_2)$ is easy; we just take the tensor product of $(\mathbf{j}_1, \mathbf{0})$ and $(\mathbf{0}, \mathbf{j}_2)$! To see this, let $v_{0,0} \in (\mathbf{j}, \mathbf{0})$ and $\bar{v}_{0,0} \in (\mathbf{0}, \mathbf{k})$ be highest weight vectors. Then it is straightforward to check that

$$v_{0,0} \otimes \bar{v}_{0,0} \in (\mathbf{j}, \mathbf{0}) \otimes (\mathbf{0}, \mathbf{k}) \tag{5.92}$$

is a highest weight vector with $j_1 = j$, $j_2 = k$, and so we conclude that $(\mathbf{j}, \mathbf{0}) \otimes (\mathbf{0}, \mathbf{k})$ contains an invariant subspace equivalent to (\mathbf{j}, \mathbf{k}) . However, since

$$\dim[(\mathbf{j}, \mathbf{0}) \otimes (\mathbf{0}, \mathbf{k})] = \dim(\mathbf{j}, \mathbf{k}) = (2j+1)(2k+1)$$

we conclude that these representations are equivalent, and so in general (switching notation a little),

$$\boxed{(\mathbf{j}_1, \mathbf{j}_2) \simeq (\mathbf{j}_1, \mathbf{0}) \otimes (\mathbf{0}, \mathbf{j}_2)}. \tag{5.93}$$

Thus, **all the irreps of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ can be built out of the left-handed spinor (fundamental) representation, the right-handed spinor representation, and various tensor products of the two.**

As before, this classification of the complex finite-dimensional irreps of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ also yields, with minimal effort, the classification of the complex finite-dimensional irreps of $SL(2, \mathbb{C})$. Any complex finite-dimensional irrep of $SL(2, \mathbb{C})$ yields a complex finite-dimensional irrep of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, which must be equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$ for some j_1, j_2 and since

$$(\mathbf{j}_1, \mathbf{j}_2) \simeq (S^{2j_1} \pi \otimes S^{2j_2} \bar{\pi}, S^{2j_1}(\mathbb{C}^2) \otimes S^{2j_2}(\mathbb{C}^2))$$

we conclude that our original $SL(2, \mathbb{C})$ irrep is equivalent to $(S^{2j_1} \Pi \otimes S^{2j_2} \bar{\Pi}, S^{2j_1}(\mathbb{C}^2) \otimes S^{2j_2}(\mathbb{C}^2))$ for some j_1, j_2 . Thus, the representations

$$(S^{2j_1} \Pi \otimes S^{2j_2} \bar{\Pi}, S^{2j_1}(\mathbb{C}^2) \otimes S^{2j_2}(\mathbb{C}^2)), \quad 2j_1, 2j_2 \in \mathbb{Z}$$

are (up to equivalence) all the complex finite-dimensional irreducible representations of $SL(2, \mathbb{C})$.

How about representations of $SO(3, 1)_o$? We saw that in the case of $SO(3)$, not all representations of the associated Lie algebra actually arise from representations of the group, and the same is true here. Say we have a complex finite-dimensional irrep (Π, V) of $SO(3, 1)_o$, and consider its induced Lie algebra representation (π, V) , which must be equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$ for some j_1, j_2 . Noting that

$$iM_z + iN_z = i\pi(S_z)$$

we have (again, be sure to distinguish π the number from π the representation!)

$$e^{i2\pi \cdot (iM_z + iN_z)} v_{0,0} = e^{i2\pi(j_1 + j_2)} v_{0,0}$$

as well as

$$\begin{aligned} e^{i2\pi \cdot (iM_z + iN_z)} v_{0,0} &= e^{-2\pi \cdot \pi(S_z)} v_{0,0} \\ &= \Pi(e^{-2\pi S_z}) v_{0,0} \\ &= \Pi(I) v_{0,0} \\ &= v_{0,0} \end{aligned}$$

so we conclude that

$$e^{i2\pi(j_1 + j_2)} = 1 \iff j_1 + j_2 \in \mathbb{Z}, \quad (5.94)$$

and thus only representations $(\mathbf{j}_1, \mathbf{j}_2)$ satisfying this condition can arise from $SO(3, 1)_o$ representations. (It is also true that for any j_1, j_2 satisfying this condition, there exists an $SO(3, 1)_o$ representation with induced Lie algebra representation $(\mathbf{j}_1, \mathbf{j}_2)$, though we will not prove that here.)

Example 5.38 \mathbb{R}^4 : *The four-vector representation of $SO(3, 1)_o$*

The fundamental representation (Π, \mathbb{R}^4) is the most familiar $SO(3, 1)_o$ representation, corresponding to four-dimensional vectors in Minkowski space. What

$(\mathbf{j}_1, \mathbf{j}_2)$ does it correspond to? To find out, we first complexify the representation to $(\Pi_{\mathbb{C}}, \mathbb{C}^4)$ and then consider the induced $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representation $(\pi_{\mathbb{C}}, \mathbb{C}^4)$. Straightforward calculations show that

$$iM_z = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$iN_z = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

and this, along with expressions for M_+ and N_+ that you should derive, can be used to show that

$$v_{0,0} = (e_1 + ie_2) = \begin{pmatrix} 1 \\ i \\ 0 \\ 0 \end{pmatrix}$$

is a highest weight vector with $(j_1, j_2) = (1/2, 1/2)$. Noting that

$$\dim \mathbb{C}^4 = \dim \left(\frac{1}{2}, \frac{1}{2} \right) = 4$$

we conclude that

$$\boxed{(\pi_{\mathbb{C}}, \mathbb{C}^4) \simeq \left(\frac{1}{2}, \frac{1}{2} \right)}.$$

Note that $j_1 + j_2 = 1/2 + 1/2 = 1 \in \mathbb{Z}$, in accordance with (5.94). \square

Before moving on to our next example, we need to discuss tensor products of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps. The nice thing here is that we can use what we know about the tensor product of $\mathfrak{su}(2)$ irreps to compute the decomposition of the tensor product of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps. In fact, the following is true.

Proposition 5.8 *The decomposition into irreps of the tensor product of two $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps $(\mathbf{j}_1, \mathbf{j}_2)$ and $(\mathbf{k}_1, \mathbf{k}_2)$ is given by*

$$(\mathbf{j}_1, \mathbf{j}_2) \otimes (\mathbf{k}_1, \mathbf{k}_2) = \bigoplus (\mathbf{l}_1, \mathbf{l}_2)$$

where $|j_1 - k_1| \leq l_1 \leq j_1 + k_1$, $|j_2 - k_2| \leq l_2 \leq j_2 + k_2$ (5.95)

and each $(\mathbf{l}_1, \mathbf{l}_2)$ consistent with the above inequalities occurs exactly once in the direct sum decomposition.

Notice the restrictions on l_1 and l_2 , which correspond to the decomposition of tensor products of $\mathfrak{su}(2)$ representations. We relegate a proof of this formula to the

Appendix, but it should seem plausible. We have seen that one can roughly think of an $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representations as a ‘product’ of two $\mathfrak{su}(2)$ representations, and so the tensor product of two $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representations (which can both be ‘factored’ into $\mathfrak{su}(2)$ representations) should just be given by the various ‘products’ of $\mathfrak{su}(2)$ representations that occur when taking the tensor product of the factors. We will apply this formula and make this concrete in the next example.

Example 5.39 $\Lambda^2 \mathbb{R}^4$: *The antisymmetric tensor representation of $SO(3, 1)_o$*

This is an important example since the electromagnetic field tensor $F^{\mu\nu}$ lives in this representation. To classify this representation, we first note that $\Lambda^2 \mathbb{R}^4$ occurs in the $O(3, 1)$ -invariant decomposition

$$\mathbb{R}^4 \otimes \mathbb{R}^4 = \mathbb{R} \eta^{-1} \oplus \Lambda^2(\mathbb{R}^4) \oplus S^{2'}(\mathbb{R}^4) \quad (5.96)$$

where

$$\eta^{-1} = \eta^{\mu\nu} e_{\mu} \otimes e_{\nu}$$

is the inverse of the Minkowski metric and

$$S^{2'}(\mathbb{R}^4) = \{T \in S^2(\mathbb{R}^4) \mid \eta_{\mu\nu} T^{\mu\nu} = 0\} \quad (5.97)$$

is the set of symmetric ‘traceless’ second rank tensors, where the trace is effected by the Minkowski metric η . Note that this is just the $O(3, 1)$ analog of the $O(n)$ decomposition in Example 5.23. You should check that each of the subspaces in (5.96) really is $O(3, 1)$ invariant. Complexifying this yields

$$\mathbb{C}^4 \otimes \mathbb{C}^4 = \mathbb{C} \eta^{-1} \oplus \Lambda^2(\mathbb{C}^4) \oplus S^{2'}(\mathbb{C}^4).$$

Now, we can also decompose $\mathbb{C}^4 \otimes \mathbb{C}^4$ using Proposition 5.8, which yields

$$\left(\frac{1}{2}, \frac{1}{2}\right) \otimes \left(\frac{1}{2}, \frac{1}{2}\right) = (\mathbf{0}, \mathbf{0}) \oplus (\mathbf{1}, \mathbf{0}) \oplus (\mathbf{0}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{1}). \quad (5.98)$$

Now, clearly $\mathbb{C} \eta^{-1}$ corresponds to $(\mathbf{0}, \mathbf{0})$ since the former is a one-dimensional representation and $(\mathbf{0}, \mathbf{0})$ is the only one-dimensional irrep in the decomposition (5.98). What about $S^{2'}(\mathbb{C}^4)$? Well, it is straightforward to check using the results of the previous example that

$$v_{0,0} = (e_1 + ie_2) \otimes (e_1 + ie_2) \in S^{2'}(\mathbb{C}^4)$$

is a highest weight vector with $(j_1, j_2) = (1, 1)$ (it is also instructive to verify that $v_{0,0}$ actually satisfies the condition in (5.97)). Checking dimensions then tells us that $S^{2'}(\mathbb{C}^4) \simeq (\mathbf{1}, \mathbf{1})$, so we conclude that

$$\boxed{\Lambda^2(\mathbb{C}^4) \simeq (\mathbf{1}, \mathbf{0}) \oplus (\mathbf{0}, \mathbf{1}).}$$

This representation is decomposable, but remember that this does *not* imply that $\Lambda^2(\mathbb{R}^4)$ is decomposable! In fact, $\Lambda^2(\mathbb{R}^4)$ is irreducible. This is an unavoidable

subtlety of the relationship between complex representations and real representations.¹⁶ For an interpretation of the representations $(\mathbf{1}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{1})$ individually, see Problem 5.12.

5.11 Irreducibility and the Representations of $O(3, 1)$ and Its Double Covers

In this section we will examine the constraints that parity and time-reversal place on representations of $O(3, 1)$ and its double covers. In particular, we will clarify our discussion of the Dirac spinor from Example 5.26 and explain why such $\mathfrak{so}(3, 1)$ decomposable representations seem to occur so naturally.

To start, consider the adjoint representation $(\text{Ad}, \mathfrak{so}(3, 1))$ of $O(3, 1)$. It is easily checked that the parity operator acts as

$$\begin{aligned}\text{Ad}_P(\tilde{L}_i) &= \tilde{L}_i \\ \text{Ad}_P(K_i) &= -K_i.\end{aligned}$$

Now say that we have a double cover of $O(3, 1)$, call it H (these certainly exist and are actually non-unique; see Sternberg [16] and the comments at the end of this section). H will have multiple components, just as $O(3, 1)$ does, and the component containing the identity will be isomorphic to $SL(2, \mathbb{C})$.¹⁷ Since H is a double cover, there exists a two-to-one group homomorphism $\Phi : H \rightarrow O(3, 1)$, which induces the usual Lie algebra isomorphism $\phi : \mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}} \rightarrow \mathfrak{so}(3, 1)$. Now let $\tilde{P} \in H$ cover $P \in O(3, 1)$, so that $\Phi(\tilde{P}) = P$. Then from the identity

$$\phi(\text{Ad}_h(X)) = \text{Ad}_{\Phi(h)}(\phi(X)) \quad \forall h \in H, X \in \mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}} \quad (5.99)$$

which you will prove below, we have

$$\begin{aligned}\phi(\text{Ad}_{\tilde{P}}(S_i)) &= \text{Ad}_P(\tilde{L}_i) \\ &= \tilde{L}_i\end{aligned}$$

as well as

$$\phi(\text{Ad}_{\tilde{P}}(\tilde{K}_i)) = -K_i.$$

Since ϕ is an isomorphism we conclude that

$$\begin{aligned}\text{Ad}_{\tilde{P}}(S_i) &= \tilde{P} S_i \tilde{P}^{-1} = S_i \\ \text{Ad}_{\tilde{P}}(\tilde{K}_i) &= \tilde{P} \tilde{K}_i \tilde{P}^{-1} = -\tilde{K}_i.\end{aligned}$$

If (Π, V) is a representation of H and (π, V) the induced $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representation, then this implies that

¹⁶For the whole story on this relationship, see Onischik [11].

¹⁷This should seem plausible, but proving it rigorously would require homotopy theory and would take us too far afield. See Frankel [4] for a nice discussion of this topic.

$$\Pi_{\tilde{P}} M_i \Pi_{\tilde{P}}^{-1} = N_i \quad (5.100)$$

$$\Pi_{\tilde{P}} N_i \Pi_{\tilde{P}}^{-1} = M_i. \quad (5.101)$$

Let us examine some consequences of this. Let $W \subset V$ be an irreducible subspace of (π, V) equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$, spanned by our usual basis of the form

$$\mathcal{B} = \{v_{k_1, k_2} \mid 0 \leq k_1 \leq 2j_1, 0 \leq k_2 \leq 2j_2\}.$$

We then have

$$iM_z \Pi_{\tilde{P}} v_{0,0} = i \Pi_{\tilde{P}} N_z v_{0,0} = j_2 \Pi_{\tilde{P}} v_{0,0}$$

$$iN_z \Pi_{\tilde{P}} v_{0,0} = i \Pi_{\tilde{P}} M_z v_{0,0} = j_1 \Pi_{\tilde{P}} v_{0,0}$$

$$M_+ \Pi_{\tilde{P}} v_{0,0} = \Pi_{\tilde{P}} N_+ v_{0,0} = 0$$

$$N_+ \Pi_{\tilde{P}} v_{0,0} = \Pi_{\tilde{P}} M_+ v_{0,0} = 0$$

and thus $\Pi_{\tilde{P}} v_{0,0}$ is a highest weight vector for $(\mathbf{j}_2, \mathbf{j}_1)$! We have thus proven the following proposition:

Proposition 5.9 *Let H be a double cover of $O(3, 1)$ and (Π, V) a complex representation of H with induced $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representation (π, V) . If $W \subset V$ is an irreducible subspace of (π, V) equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$ and $j_1 \neq j_2$, then there exists another irreducible subspace W' of (π, V) equivalent to $(\mathbf{j}_2, \mathbf{j}_1)$.*

It should be clear from the above that the operator $\Pi(\tilde{P})$ corresponding to parity takes us back and forth between W and W' . This means that even though W is an invariant subspace of the Lie algebra representation (π, V) , W is *not* invariant under the Lie group representation (Π, V) , since $\Pi(\tilde{P})$ takes vectors in W to vectors in W' ! If W and W' make up all of V , i.e. if $V = W \oplus W'$, this means that V is irreducible under the H representation Π but not under the $\mathfrak{so}(3, 1)$ representation π , and so we have an irreducible Lie group representation whose induced Lie algebra representation is *not* irreducible! We will meet two examples of this type of representation below. Note that this does not contradict Proposition 5.3, as the group H does not satisfy the required hypotheses of connectedness.

Exercise 5.41 Let $\Phi : H \rightarrow G$ be a Lie group homomorphism with induced Lie algebra homomorphism $\phi : \mathfrak{h} \rightarrow \mathfrak{g}$. Use the definition of the adjoint mapping and of ϕ to show that

$$\phi(\text{Ad}_h(X)) = \text{Ad}_{\Phi(h)}(\phi(X)) \quad \forall h \in H, X \in \mathfrak{h}.$$

Exercise 5.42 Verify (5.100) and (5.101). You will need the result of the previous exercise!

Example 5.40 The Dirac spinor revisited

As a particular application of Proposition 5.9, suppose our representation (Π, V) of H contains a subspace equivalent to the left-handed spinor $(\frac{1}{2}, \mathbf{0})$; then it must also contain a subspace equivalent to the right-handed spinor $(\mathbf{0}, \frac{1}{2})$. This is why the Dirac spinor is $(\frac{1}{2}, \mathbf{0}) \oplus (\mathbf{0}, \frac{1}{2})$. It is not irreducible as an $SL(2, \mathbb{C})$ representation, but it is irreducible as a representation of a group H which extends $SL(2, \mathbb{C})$ and covers $O(3, 1)$. \square

The Dirac spinor representation is a representation of H , not $O(3, 1)$, but many of the other $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}} \simeq \mathfrak{so}(3, 1)$ representations of interest do come from $O(3, 1)$ representations (for instance, the fundamental representation of $\mathfrak{so}(3, 1)$ and its various tensor products). Since any $O(3, 1)$ representation $\Pi : O(3, 1) \rightarrow GL(V)$ yields an H representation $\Pi \circ \Phi : H \rightarrow GL(V)$, Proposition 5.9 as well as the comments following it hold for $O(3, 1)$ representations as well as representations of H . In the $O(3, 1)$ case, though, we can do even better:

Proposition 5.10 *Let (Π, V) be a finite-dimensional complex irreducible representation of $O(3, 1)$. Then the induced $\mathfrak{so}(3, 1)$ representation (π, V) is equivalent to one of the following:*

$$(\mathbf{j}, \mathbf{j}), \quad 2j \in \mathbb{Z} \quad \text{or} \quad (\mathbf{j}_1, \mathbf{j}_2) \oplus (\mathbf{j}_2, \mathbf{j}_1), \quad 2j_1, 2j_2 \in \mathbb{Z}, \quad j_1 + j_2 \in \mathbb{Z}, \quad j_1 \neq j_2. \quad (5.102)$$

Proof Since $\mathfrak{so}(3, 1)$ is semi-simple, (π, V) is completely reducible, i.e. equivalent to a direct sum of irreducible representations. Let $W \subset V$ be one such representation, equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$, with highest weight vector $v_{0,0}$. Then $\Pi(P)v_{0,0} \equiv v_P$ is a highest weight vector with eigenvalues (j_2, j_1) , and the same arguments show that $\Pi(T)v_{0,0} \equiv v_T$ is also a highest weight vector with eigenvalues (j_2, j_1) (it is not necessarily equal to v_P , though). The same arguments also show that $\Pi(PT)v_{0,0} \equiv v_{PT}$ is a highest weight vector with eigenvalues (j_1, j_2) (recall that T is the time-reversal operator defined in (4.34)). Now consider the vectors

$$\begin{aligned} w_{0,0} &\equiv v_{0,0} + v_{PT} \\ u_{0,0} &\equiv v_P + v_T. \end{aligned}$$

$w_{0,0}$ is clearly a highest weight vector with eigenvalues (j_1, j_2) , and $u_{0,0}$ is clearly a highest weight vector with eigenvalues (j_2, j_1) . Using the fact that P and T commute, you can easily check that

$$\begin{aligned} \Pi(P)w_{0,0} &= \Pi(T)w_{0,0} = u_{0,0} \\ \Pi(P)u_{0,0} &= \Pi(T)u_{0,0} = w_{0,0}. \end{aligned}$$

We can then define basis vectors

$$\begin{aligned} w_{k_1, k_2} &\equiv (M_-)^{k_1} (N_-)^{k_2} w_{0,0}, \quad 0 \leq k_1 \leq 2j_1, \quad 0 \leq k_2 \leq 2j_2 \\ u_{k_1, k_2} &\equiv (M_-)^{k_1} (N_-)^{k_2} u_{0,0}, \quad 0 \leq k_1 \leq 2j_2, \quad 0 \leq k_2 \leq 2j_1 \end{aligned}$$

which span $\mathfrak{so}(3, 1)$ irreducible subspaces $W \simeq (\mathbf{j}_1, \mathbf{j}_2)$ and $U \simeq (\mathbf{j}_2, \mathbf{j}_1)$. By Proposition 5.3 and the connectedness of $SO(3, 1)_o$, W and U are also irreducible under $SO(3, 1)_o$, and from (5.94) we know that $j_1 + j_2 \in \mathbb{Z}$. Furthermore, by the definition of the w_{k_1, k_2} and u_{k_1, k_2} , as well as (5.100)–(5.101), we have

$$\begin{aligned} \Pi(P)w_{k_1, k_2} &= \Pi(T)w_{k_1, k_2} = u_{k_2, k_1} \\ \Pi(P)u_{k_1, k_2} &= \Pi(T)u_{k_1, k_2} = w_{k_2, k_1} \end{aligned}$$

and so $\text{Span}\{w_{k_1, k_2}, u_{l_1, l_2}\}$ is invariant under $O(3, 1)$. We assumed V was irreducible, though, so we conclude that $V = \text{Span}\{w_{k_1, k_2}, u_{l_1, l_2}\}$. Does that mean we can conclude that $V \simeq (\mathbf{j}_1, \mathbf{j}_2) \oplus (\mathbf{j}_2, \mathbf{j}_1)$? Not quite, because we never established that $w_{0,0}$ and $u_{0,0}$ were linearly independent! In fact, they might be linearly dependent, in which case they would be proportional, which would imply that each w_{k_1, k_2} is proportional to u_{k_1, k_2} (why?), and also that $j_1 = j_2 \equiv j$. In this case, we obtain

$$V = \text{Span}\{w_{k_1, k_2}, u_{l_1, l_2}\} = \text{Span}\{w_{k_1, k_2}\} \simeq (\mathbf{j}, \mathbf{j})$$

which is one of the alternatives mentioned in the statement of the proposition. If $w_{0,0}$ and $u_{0,0}$ are linearly independent, however, then so is the set $\{w_{k_1, k_2}, u_{l_1, l_2}\}$ and so

$$V = \text{Span}\{w_{k_1, k_2}, u_{l_1, l_2}\} = W \oplus U \simeq (\mathbf{j}_1, \mathbf{j}_2) \oplus (\mathbf{j}_2, \mathbf{j}_1)$$

which is the other alternative. All that remains is to show that $j_1 \neq j_2$, which you will do in Exercise 5.43 below. This concludes the proof. \square

Exercise 5.43 Assume that $w_{0,0}$ and $u_{0,0}$ are linearly independent and that $j_1 = j_2$. Use this to construct a non-trivial $O(3, 1)$ -invariant subspace of V , contradicting the irreducibility of V . Thus if V is irreducible and $w_{0,0}$ and $u_{0,0}$ are linearly independent, then $j_1 \neq j_2$ as desired.

Example 5.41 $O(3, 1)$: Representations revisited

In this example we just point out that all the $O(3, 1)$ representations we have met have the form (5.102). Below is a table of some of these representations, along with their complexifications and the corresponding $\mathfrak{so}(3, 1)$ representations.

Name	V	$V_{\mathbb{C}}$	$\mathfrak{so}(3, 1)$ rep
Scalar (trivial)	\mathbb{R}	\mathbb{C}	$(\mathbf{0}, \mathbf{0})$
Vector (fundamental)	\mathbb{R}^4	\mathbb{C}^4	$(\frac{1}{2}, \frac{1}{2})$
Antisymmetric Tensor (adjoint)	$\Lambda^2 \mathbb{R}^4$	$\Lambda^2 \mathbb{C}^4$	$(\mathbf{1}, \mathbf{0}) \oplus (\mathbf{0}, \mathbf{1})$
Pseudovector	$\Lambda^3 \mathbb{R}^4$	$\Lambda^3 \mathbb{C}^4$	$(\frac{1}{2}, \frac{1}{2})$
Pseudoscalar	$\Lambda^4 \mathbb{R}^4$	$\Lambda^4 \mathbb{C}^4$	$(\mathbf{0}, \mathbf{0})$
Symmetric Traceless Tensor	$S^{2'}(\mathbb{R}^4)$	$S^{2'}(\mathbb{C}^4)$	$(\mathbf{1}, \mathbf{1})$

Note that the pseudovector and pseudoscalar representations yield the same $\mathfrak{so}(3, 1)$ representations as the vector and scalar, respectively, as discussed in Example 5.18. Thus we have a pair of examples in which two equivalent Lie algebra representations come from two *non*-equivalent matrix Lie group representations! Again, this does not contradict Proposition 5.2 since $O(3, 1)$ is not connected. Note also that the only representation in the above table that decomposes into more than one $\mathfrak{so}(3, 1)$ irrep is the antisymmetric tensor representation; see Problem 5.12 for

the action of the parity operator on this representation, and how it takes one back and forth between the $\mathfrak{so}(3, 1)$ -irreducible subspaces $(\mathbf{1}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{1})$. \square

Before concluding this chapter we should talk a little bit about this mysterious group H which is supposed to be a double cover of $O(3, 1)$. It can be shown¹⁸ that there are exactly eight non-isomorphic double covers of $O(3, 1)$ (in contrast to the case of $SO(3)$ and $SO(3, 1)_o$ which have the unique connected double covers $SU(2)$ and $SL(2, \mathbb{C})$). Most of these double covers are somewhat obscure and do not really crop up in the physics literature, but two of them are quite natural and well-studied: these are the **Pin** groups $\text{Pin}(3, 1)$ and $\text{Pin}(1, 3)$ which appear in the study of **Clifford Algebras**. Clifford algebras are rich and beautiful objects, and lead naturally to double covers of all the orthogonal and Lorentz groups. In the four-dimensional Lorentzian case in particular, one encounters the Dirac gamma matrices and the Dirac spinor, as well as the Pin groups which act naturally on the Dirac spinor. For details on the construction of the Pin groups and their properties, see Göckeler and Schücker [7].

5.12 Problems

5.1 In this problem we will develop a coordinate-based proof of our claim from Example 5.22.

(a) Let $X = X^{ij} e_i \otimes e_j \in \Lambda^2 \mathbb{R}^n$, and define

$$\begin{aligned}\tilde{X} &\equiv X^{ij} L(e_i) \otimes e_j = X^{ij} g_{ik} e^k \otimes e_j \in \mathcal{L}(\mathbb{R}^n) \\ \phi(X) &\equiv [\tilde{X}] \in M_n(\mathbb{R}).\end{aligned}$$

Find an expression for $\phi(X)$ in terms of $[X]$ and use it to show that $\phi(X) \in \mathfrak{g}$.

(b) Prove that $\text{Ad}(R) \circ \phi = \phi \circ \Lambda^2 \Pi(R)$ by evaluating both sides on an arbitrary $X \in \Lambda^2 \mathbb{R}^n$ and showing that the components of the matrices are equal. You will need the expansion of X given above, as well as the coordinate form (5.43) of $\Lambda^2 \Pi$. For simplicity of matrix computation, you may wish to abandon the Einstein Summation Convention here and write the components of $R \in G$ as R_{ij} , even though you are interpreting R as a linear operator.

5.2 In this problem we will develop a coordinate-free proof that the fundamental representation of $SL(2, \mathbb{C})$ is equivalent to its dual. This will also imply that the fundamental representation of $SU(2)$ is equivalent to its dual as well.

(a) Consider the epsilon tensor in $\mathcal{T}_0^2(\mathbb{C}^2)$,

$$\epsilon \equiv e^1 \wedge e^2.$$

Using the definition (3.84) of the determinant, show that ϵ is $SL(2, \mathbb{C})$ invariant, i.e. that

¹⁸See Sternberg [16].

$$\epsilon(Av, Aw) = \epsilon(v, w) \quad \forall v, w \in \mathbb{C}^2, \quad A \in SL(2, \mathbb{C}).$$

(b) Define a map

$$L_\epsilon : \mathbb{C}^2 \rightarrow \mathbb{C}^{2*}$$

$$v \mapsto \epsilon(v, \cdot).$$

Using your result from (a), show that L_ϵ is an intertwiner, and also show that L_ϵ is one-to-one. Conclude that $\mathbb{C}^2 \simeq \mathbb{C}^{2*}$ as $SL(2, \mathbb{C})$ irreps, and hence as $SU(2)$ irreps as well.

(c) Assuming the very plausible but slightly annoying to prove fact that $(S^{2l}(\mathbb{C}^2))^* \simeq S^{2l}(\mathbb{C}^{2*})$, prove that *every* $SU(2)$ representation is equivalent to its dual. One might think that this is self-evident since taking the dual of an irrep does not change its dimension and for a given dimension there is only one $SU(2)$ irrep (up to equivalence), but this assumes that the dual of an irrep is itself an irrep. This is true, but needs to be proven, and is the subject of the next problem.

5.3 Show that if (Π, V) is a finite-dimensional irreducible group or Lie algebra representation, then so is (Π^*, V^*) . Note that this together with Problem 2.4 also implies the converse, namely that if (Π^*, V^*) is an irrep then so is (Π, V) . Use this to again prove that every $\mathfrak{su}(2)$ irrep is equivalent to its dual.

5.4 In this problem we will prove in coordinates that the fundamental of $\mathfrak{su}(2)$ is equivalent to its dual. Let $\phi : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ be an intertwiner between these two representations, so that

$$\phi \circ X = -X^T \circ \phi \quad \forall X \in \mathfrak{su}(2).$$

If we represent ϕ by a matrix M and note that by linearity it suffices to check this equation for the basis $\{S_i\}_{i=1,2,3}$ of $\mathfrak{su}(2)$, this becomes

$$MS_iM^{-1} = -S_i^T, \quad i = 1, 2, 3.$$

Find a matrix M that satisfies this equation. (Hint: Use the intertwiner from Problem 5.2.)

5.5 If G is a matrix Lie group and (Π, V) its fundamental representation, one can sometimes generate new representations by considering the *conjugate* representation

$$\bar{\Pi} : G \rightarrow GL(V)$$

$$A \mapsto \bar{A}$$

where \bar{A} denotes the matrix whose entries are just the complex conjugates of the entries of A .

(a) Verify that $\bar{\Pi}$ is a homomorphism, hence a bona fide representation. Show that for $G = SU(n)$, the conjugate representation is equivalent to the dual representation. What does this mean in the case of $SU(2)$?

- (b) Let $G = SL(2, \mathbb{C})$. Show that the conjugate to $(\frac{1}{2}, \mathbf{0})$ is $(\mathbf{0}, \frac{1}{2})$. There are a few ways to do this. This justifies the notation $\bar{\Pi}$ for the representation homomorphism of $(\mathbf{0}, \frac{1}{2})$.

5.6 Let (π, V) be a representation of a Lie algebra \mathfrak{g} and assume that $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is one-to-one (such Lie algebra representations are said to be *faithful*). Show that there is an invariant subspace of $(\pi_1^1, V \otimes V^*)$ that is equivalent to the adjoint representation $(\text{ad}, \mathfrak{g})$.

5.7 In this problem we will meet a representation that is *not* completely reducible.

- (a) Consider the representation (Π, \mathbb{R}^2) of \mathbb{R} given by

$$\begin{aligned}\Pi : \mathbb{R} &\rightarrow GL(2, \mathbb{R}) \\ a &\mapsto \Pi(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.\end{aligned}$$

Verify that (Π, \mathbb{R}^2) is a representation.

- (b) If Π was completely reducible then we would be able to decompose \mathbb{R}^2 as $\mathbb{R}^2 = V \oplus W$ where V and W are one-dimensional. Show that such a decomposition is impossible.

5.8 In this problem we will show that \mathbb{Z}_2 is semi-simple, i.e. that every finite-dimensional representation of \mathbb{Z}_2 is completely reducible. Our strategy will be to construct a \mathbb{Z}_2 -invariant inner product on our vector space, and then show that for any invariant subspace W , its orthogonal complement W^\perp is also invariant. We can then iterate this procedure to obtain a complete decomposition.

- (a) Let (Π, V) be a representation of \mathbb{Z}_2 . We first construct a \mathbb{Z}_2 -invariant inner product. To do this, we start with an *arbitrary* inner product $(\cdot, \cdot)_0$ on V (which could be defined, for instance, as one for which some arbitrary set of basis vectors is orthonormal). We then define a ‘group averaged’ inner product as

$$(v|w) \equiv \sum_{h \in \mathbb{Z}_2} (\Pi_h v | \Pi_h w)_0.$$

Show that $(\cdot | \cdot)$ is \mathbb{Z}_2 -invariant, i.e. that

$$(\Pi_g v | \Pi_g w) = (v|w) \quad \forall v, w \in V, \quad g \in \mathbb{Z}_2.$$

- (b) Assume now that there exists a non-trivial invariant subspace $W \subset V$ (if no such W existed, then V would be irreducible, hence completely reducible and we would be done). Define its *orthogonal complement*

$$W^\perp \equiv \{v \in V \mid (v|w) = 0 \quad \forall w \in W\}.$$

Argue that there exists an orthonormal basis

$$\mathcal{B} = \{w_1, \dots, w_k, v_{k+1}, \dots, v_n\} \quad \text{where } w_i \in W, \quad i = 1 - k,$$

and conclude that $V = W \oplus W^\perp$.

- (c) Show that W^\perp is an invariant subspace, so that $V = W \oplus W^\perp$ is a decomposition into invariant subspaces. We still do not know that W is irreducible, though. Argue that we can nonetheless iterate our above argument until we obtain a decomposition into irreducibles, thus proving that (Π, V) is completely reducible.

5.9 In this problem we will sketch a proof that the space of harmonic polynomials $\mathcal{H}_l(\mathbb{R}^3)$ has dimension $2l + 1$.

- (a) Recall our notation $P_k(\mathbb{R}^3)$ for the space of k th degree polynomials on \mathbb{R}^3 . Assume that the map

$$\Delta : P_k(\mathbb{R}^3) \longrightarrow P_{k-2}(\mathbb{R}^3) \quad (5.103)$$

is onto. Use the rank-nullity theorem to show that $\dim \mathcal{H}_l(\mathbb{R}^3) = 2l + 1$. If you need help you might consult Example 3.24, as well as (3.104).

- (b) To complete the proof we need to show that (5.103) is onto. I have not seen a clean or particularly enlightening proof of this fact, so I do not heartily recommend this part of the problem, but if you really want to show it you might try showing (inductively on k) that the two-dimensional Laplacian

$$\Delta : P_k(\mathbb{R}^2) \longrightarrow P_{k-2}(\mathbb{R}^2)$$

is onto, and then use this to show (inductively on k) that (5.103) is onto as well.

5.10 Let T and U be *commuting* linear operators on a complex vector space V , so that $[T, U] = 0$. We will show that T and U have a simultaneous eigenvector.

Using the standard argument we employed in the proof of Proposition 5.4, show that T has at least one eigenvector. Denote that vector by v_a and its eigenvalue by a . Let V_a denote the span of *all* eigenvectors of T with eigenvalue a ; V_a may just be the one-dimensional subspace spanned by v_a , or it may be bigger if there are other eigenvectors that also have eigenvalue a . Use the fact that U commutes with T to show that V_a is invariant under U (i.e. $U(v) \in V_a$ whenever $v \in V_a$). We can then restrict U to V_a to get a linear operator

$$U|_{V_a} \equiv U_a \in \mathcal{L}(V_a). \quad (5.104)$$

Then use the standard argument again to show that U_a has an eigenvector $v_0 \in V_a$. Show that this is a simultaneous eigenvector of T and U .

5.11 In this problem we would like to show that the tensor product ‘distributes’ over direct sums, in the sense that for any vector spaces V , W and Z , there exists a vector space isomorphism

$$\phi : (V \oplus W) \otimes Z \rightarrow (V \otimes Z) \oplus (W \otimes Z).$$

Define such a map on decomposable elements by

$$\phi : (v, w) \otimes z \mapsto (v \otimes z, w \otimes z)$$

and extend linearly. Show that ϕ is linear (you will have to be careful about how addition works in the direct sum spaces and the tensor product spaces) as well as one-to-one and onto, so that it is a vector space isomorphism.

5.12 In this problem we will study the $O(3, 1)$ representation $(\Lambda^2 \Pi, \Lambda^2 \mathbb{C}^4)$ and its decomposition into $(\mathbf{1}, \mathbf{0}) \oplus (\mathbf{0}, \mathbf{1})$ under $SO(3, 1)_o$ and $\mathfrak{so}(3, 1)$.

- (a) Define the *star* operator $*$ on $\Lambda^2 \mathbb{C}^4$ by

$$\begin{aligned} * : \Lambda^2 \mathbb{C}^4 &\rightarrow \Lambda^2 \mathbb{C}^4 \\ e_i \wedge e_j &\mapsto \epsilon_{ijkl} \eta^{km} \eta^{ln} e_m \wedge e_n \end{aligned} \quad (5.105)$$

and extending linearly. Here η^{ln} are the components of the inverse of the Minkowski metric and ϵ_{ijkl} are the components of the usual epsilon tensor on \mathbb{R}^4 . Show that $*$ is an intertwiner between $\Lambda^2 \mathbb{C}^4$ and itself when viewed as an $SO(3, 1)_o$ representation, but not as an $O(3, 1)$ representation. You will need (4.15) as well as your results from part (a) of Problem 3.4.

- (b) Compute the action of $*$ on the basis vectors f_i , $i = 1-6$ defined in Example 5.24. Use the f_i to construct eigenvectors of $*$, and show that $*$ is diagonalizable with eigenvalues $\pm i$, so that $\Lambda^2 \mathbb{C}^4$ decomposes into $V_{+i} \oplus V_{-i}$ where V_{+i} is the eigenspace of $*$ in which every vector is an eigenvector with eigenvalue $+i$ and likewise for V_{-i} . Then use Schur's lemma to conclude that any $SO(3, 1)_o$ -irreducible subspace must lie entirely in V_{+i} or V_{-i} . In particular, this means that $\Lambda^2 \mathbb{C}^4$ is not irreducible as an $SO(3, 1)_o$ or $\mathfrak{so}(3, 1)$ representation.
- (c) A convenient basis for V_{+i} which you may have discovered above is

$$\begin{aligned} v_{+1} &\equiv f_1 + if_4 \\ v_{+2} &\equiv f_2 + if_5 \\ v_{+3} &\equiv f_3 + if_6. \end{aligned}$$

Likewise, for V_{-i} we have the basis

$$\begin{aligned} v_{-1} &\equiv f_1 - if_4 \\ v_{-2} &\equiv f_2 - if_5 \\ v_{-3} &\equiv f_3 - if_6. \end{aligned}$$

Now consider the vectors

$$\begin{aligned} v_{0,0} &\equiv v_{+1} + iv_{+2} \\ w_{0,0} &\equiv v_{-1} + iv_{-2}. \end{aligned}$$

Show that these are highest weight vectors for $(\mathbf{0}, \mathbf{1})$ and $(\mathbf{1}, \mathbf{0})$, respectively. Then count dimensions to show that as $\mathfrak{so}(3, 1)$ representations, $V_{+i} \simeq (\mathbf{0}, \mathbf{1})$ and $V_{-i} \simeq (\mathbf{1}, \mathbf{0})$.

- (d) Show directly that $\Lambda^2 \Pi(P)(v_{+j}) = v_{-j}$, $j = 1, 2, 3$ and likewise for T , so that P and T interchange V_{+i} and V_{-i} , as expected.

Chapter 6

The Wigner–Eckart Theorem and Other Applications

In this chapter we will apply what we have learned about representations to quantum mechanics, discussing selection rules and in particular the notoriously confusing subjects of spherical tensors and the Wigner–Eckart theorem. We will then finish our discussion with a brief foray into the beautiful subject of Clifford algebras, in an attempt to better understand the Dirac spinor.

6.1 Tensor Operators, Spherical Tensors and Representation Operators

In this section we introduce the notion of a *representation operator*, which generalizes the notions of vector and tensor operators as well as spherical tensors. We will then use representation operators to derive the fundamental quantum mechanical selection rule which lays the foundation for the various selection rules one encounters in standard quantum mechanics courses.

You will recall from Sect. 3.7 that we defined a vector operator $\mathbf{B} = \{B_i\}_{i=1,2,3}$ to be a set of three linear operators that transform as vectors under the adjoint action of the total angular momentum operators J_i . With the machinery we have developed, we can now recast this definition a bit. More precisely, we can now say that if we have an $SO(3)$ representation (Π, \mathcal{H}) on some Hilbert space \mathcal{H} , then a **vector operator** on \mathcal{H} is a set $\mathbf{B} = \{B_i\}_{i=1,2,3}$ of three operators that satisfy

$$\Pi(R)B_i\Pi(R)^{-1} = \sum_{k=1}^3 R_{ki}B_k \quad \forall i = 1, 2, 3, \quad R \in SO(3). \quad (6.1)$$

Notice that $\Pi(R)B_i\Pi(R) = \Pi_1^1(R)B_i$, so (6.1) just says that under the induced representation Π_1^1 of $SO(3)$ on $\mathcal{L}(\mathcal{H})$, the B_i transform like the basis vectors e_i of the fundamental (vector) representation of $SO(3)$. This is still roughly how vector operators are defined in most physics texts. We can go further, however, and interpret (6.1) as saying that $\text{Span}\{B_i\}$, considered as a vector subspace of $\mathcal{L}(\mathcal{H})$, is an

$SO(3)$ -invariant subspace (under the representation Π_1^1) that is actually equivalent to the vector representation of $SO(3)$! This is an important shift in perspective, as **we are now thinking of the space of operators as a representation space, and of the B_i as spanning an invariant subspace of this space**. This point of view has some advantages, and is the one that will be formalized shortly in the notion of a representation operator. Now, to verify that the definition (6.1) is equivalent to our earlier definition of a vector operator from Chap. 3, you should check that if we perform our usual trick and set $R = e^{tL_j}$ and differentiate at $t = 0$, as well as identify $i\pi(L_i)$ with what we have called the total angular momentum operator J_i , we obtain

$$[J_i, B_j] = i \sum_{k=1}^3 \epsilon_{ijk} B_k \quad (6.2)$$

which is just (3.70).

Now, in physics texts one usually defines **tensor operators** similarly, i.e. as a set of operators $\mathbf{T}_{(s)}^{(r)} = \{T^{i_1 \dots i_r}_{j_1 \dots j_s}\}$ that under the representation Π_1^1 of $SO(3)$ transform like the basis vectors $e^{i_1} \otimes \dots \otimes e^{i_r} \otimes e_{j_1} \otimes \dots \otimes e_{j_s}$ of the space $\mathcal{T}_s^r(\mathbb{R}^3)$. Again, this can be alternately formulated as saying that $\text{Span}\{T^{i_1 \dots i_r}_{j_1 \dots j_s}\} \subset \mathcal{L}(\mathcal{H})$ is an invariant subspace equivalent to the $\mathcal{T}_s^r(\mathbb{R}^3)$ representation of $SO(3)$. We have seen, though, that tensor product spaces such as $\mathcal{T}_s^r(\mathbb{R}^3)$ are usually decomposable, so it is sometimes advantageous to consider **spherical tensors** of degree l , which are sets of $2l + 1$ operators $\mathbf{T}^{(l)} = \{T_k^{(l)}\}_{0 \leq k \leq 2l}$ that under Π_1^1 transform as the basis vectors v_k of the *irreducible* representation space V_l , $l \in \mathbb{Z}$.¹ Again, we will find it useful to think about $\text{Span}\{T_k^{(l)}\}$ as an invariant subspace of $\mathcal{L}(\mathcal{H})$ equivalent to V_l , rather than focusing on how the individual elements $T_k^{(l)}$ transform. Note that the ‘ l ’ in the superscript of $T_k^{(l)}$ is in parentheses; this is because the l is not really an active index, but just serves to remind us which $SO(3)$ representation we are dealing with. (As an additional side note, the term ‘spherical tensor’ derives from the fact that the representation V_l , $l \in \mathbb{Z}$ is equivalent to the space $\tilde{\mathcal{H}}_l$ of spherical harmonics of degree l .)

The following example will help make the foregoing concrete.

Example 6.1 Product of vector operators

Consider two vector operators $\mathbf{A} = \{A_i\}_{i=1,2,3}$ and $\mathbf{B} = \{B_j\}_{j=1,2,3}$ on some Hilbert space \mathcal{H} with $SO(3)$ representation Π . We can form a tensor operator (of type $(0, 2)$) by considering the set of their products $\mathbf{AB} = \{A_i B_j\}$, $i, j = 1, 2, 3$. Each $A_i B_j$ is a linear operator on \mathcal{H} , and it is easy to verify that \mathbf{AB} is a bona fide $(0, 2)$ tensor operator:

¹Warning: In other texts, different normalization and indexing conventions are used for the specific basis vectors whose transformation properties the spherical tensors are supposed to mimic. In most quantum mechanics texts, for instance, the components of a spherical tensor are supposed to mimic the transformation properties of the kets $|l, m\rangle$ which have different labeling and normalization conventions than our v_k , even though they are essentially the same thing.

$$\begin{aligned}
\Pi(R)A_iB_j\Pi(R)^{-1} &= [\Pi(R)A_i\Pi(R)^{-1}][\Pi(R)B_j\Pi(R)^{-1}] \\
&= \sum_{k,l=1}^3 R_{ki}R_{lj}A_kB_l
\end{aligned} \tag{6.3}$$

which is the same as the usual transformation law for the basis vectors $\{e_i \otimes e_j\}$ of $\mathcal{T}_2^0(\mathbb{R}^3)$. Thus, the nine-dimensional space $\text{Span}\{A_iB_j\}$ is equivalent to $\mathcal{T}_2^0(\mathbb{R}^3)$ as an $SO(3)$ representation! We know, however, that $\mathcal{T}_2^0(\mathbb{R}^3)$ is reducible, and so then $(\Pi_1^1, \text{Span}\{A_iB_j\})$ must be reducible as well. We can decompose $\text{Span}\{A_iB_j\}$ into irreducible subspaces just as we decomposed $\mathcal{T}_2^0(\mathbb{R}^3) = \mathbb{R}g \oplus S^{2'}(\mathbb{R}^3) \oplus \Lambda^2(\mathbb{R}^3)$, and we should correspondingly obtain spherical tensors of degrees 0, 1 and 2. Explicitly, the degree 0 subspace corresponding to $g = \sum_{i=1}^3 e_i \otimes e_i$ yields a spherical tensor $\mathbf{T}^{(0)}$ with single element

$$T_0^{(0)} = \sum_{i=1}^3 A_iB_i$$

which transforms as a scalar. The degree 1 spherical tensor $\mathbf{T}^{(1)}$ corresponding to $\Lambda^2\mathbb{R}^3$ has elements²

$$\begin{aligned}
T_0^{(1)} &= (\mathbf{A} \times \mathbf{B})_x + i(\mathbf{A} \times \mathbf{B})_y \\
T_1^{(1)} &= -2(\mathbf{A} \times \mathbf{B})_z \\
T_2^{(1)} &= -2[(\mathbf{A} \times \mathbf{B})_x - i(\mathbf{A} \times \mathbf{B})_y]
\end{aligned}$$

where notation like $(\mathbf{A} \times \mathbf{B})_x$ is just shorthand for $A_yB_z - A_zB_y$. The appearance of the cross product here shouldn't be surprising, since the cross product is essentially just a way of identifying a bivector (or antisymmetric tensor) in $\Lambda^2\mathbb{R}^3$ with a vector in \mathbb{R}^3 , cf. Example 3.29. The degree 2 spherical tensor $\mathbf{T}^{(2)}$ corresponding to $S^{2'}(\mathbb{R}^3)$ has elements

$$\begin{aligned}
T_0^{(2)} &= (A_x + iA_y)(B_x + iB_y) \\
T_1^{(2)} &= -2A_z(B_x + iB_y) - 2(A_x + iA_y)B_z \\
T_2^{(2)} &= -4(A_xB_x + A_yB_y - 2A_zB_z) \\
T_3^{(2)} &= 12A_z(B_x + iB_y) + 12(A_x + iA_y)B_z \\
T_4^{(2)} &= 24(A_x - iA_y)(B_x - iB_y).
\end{aligned}$$

To convince yourself that these operators really do transform as claimed, it is useful to compare the above expressions with the expressions for the v_k for $\mathcal{H}_1(\mathbb{R}^3)$ and $\mathcal{H}_2(\mathbb{R}^3)$ (you should have computed these earlier as the f_k from Problem 2.2). You should also convince yourself that the span of $\{A_iB_j\}$ is equal to the span

²The funny looking normalization factors and index convention come from our normalization and indexing convention for the v_k s, which, as noted above, differ from the conventions in the physics literature.

of $\mathbf{T}^{(0)}$, $\mathbf{T}^{(1)}$, and $\mathbf{T}^{(2)}$ taken together, so that we have the decomposition of $(\Pi_1^1, \text{Span}\{A_i B_j\})$ into irreducibles as

$$\text{Span}\{A_i B_j\} = \text{Span } \mathbf{T}^{(0)} \oplus \text{Span } \mathbf{T}^{(1)} \oplus \text{Span } \mathbf{T}^{(2)}.$$

One might also call this a “decomposition of $\{A_i B_j\}$ into spherical tensors”. \square

By now you may have noticed that the above definitions are somewhat unwieldy; they rely on transformation properties of specific operators, and seem to rely too heavily on those operators, since the *span* of those operators seems more significant than the operators themselves. To remedy this, we present the following definition, which subsumes all of the foregoing definitions and also presents things in a more basis-independent fashion: given a representation (Π_0, V_0) of a group G on some auxiliary vector space V_0 , as well as a unitary representation (Π, \mathcal{H}) of G on some Hilbert space \mathcal{H} , we define a **representation operator** to simply be an intertwiner between V_0 and $\mathcal{L}(\mathcal{H})$, or in other words a linear map $\rho : V_0 \rightarrow \mathcal{L}(\mathcal{H})$ satisfying

$$\boxed{\rho(\Pi_0(g)v) = \Pi(g)\rho(v)\Pi(g)^{-1} \quad \forall g \in G, v \in V_0.} \quad (6.4)$$

Note that in terms of maps between V_0 and $\mathcal{L}(\mathcal{H})$, this just says

$$\rho \circ \Pi_0(g) = \Pi_1^1(g) \circ \rho \quad \forall g \in G. \quad (6.5)$$

What this definition is saying, roughly, is that we have a subspace $\rho(V_0) \subset \mathcal{L}(\mathcal{H})$ which, even though it is composed of operators acting on \mathcal{H} , actually transforms like the space V_0 under similarity transformations by the operators Π_g . Note that, strictly speaking, the representation operator ρ is not itself an operator, but an intertwiner between representations.

How does this definition subsume the previous ones? Let V_0 have a basis $\{e_i\}$; then since ρ is linear, it is completely determined by its action on this basis, and plugging a basis vector e_i into (6.4) then yields

$$\Pi(g)\rho(e_i)\Pi(g)^{-1} = (\Pi_0(g))_i^j \rho(e_j). \quad (6.6)$$

If we set $\rho(e_i) \equiv B_i$, this becomes

$$\Pi(g)B_i\Pi(g)^{-1} = (\Pi_0(g))_i^j B_j. \quad (6.7)$$

This, of course, just says that under the representation Π_1^1 , the B_i transform like basis vectors of the representation (Π_0, V_0) , which was how we defined vector operators, tensor operators, and spherical tensors in the first place! In fact, if we take $G = SO(3)$ and $V_0 = \mathbb{R}^3$, then (6.7) becomes (writing the matrices of the fundamental representation with both indices down)

$$\Pi(R)B_i\Pi(R)^{-1} = \sum_{j=1}^3 R_{ji} B_j$$

which is of course just (6.1). To get tensor operators or spherical operators, we of course just have to take $V_0 = \mathcal{T}_s^r(\mathbb{R}^3)$ or $V_0 = \mathcal{H}_l(\mathbb{R}^3)$, respectively.

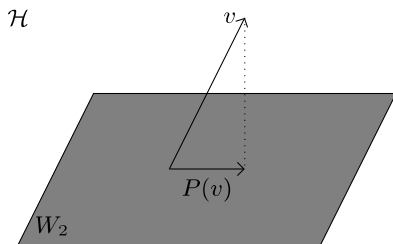


Fig. 6.1 Action of the orthogonal projection operator $P : \mathcal{H} \rightarrow W_2$ on a vector $v \in \mathcal{H}$

Before leaving this section, it is worth re-emphasizing the shift in viewpoint that we have made: we have taken a representation (Π, \mathcal{H}) of G on a Hilbert space \mathcal{H} and then used the induced representation $(\Pi_1^1, \mathcal{L}(\mathcal{H}))$ to let G act on the space of operators $\mathcal{L}(V)$; this space (of operators!) may then break up into irreducible components, and knowing what irrep an operator lives in may actually tell us quite a bit about its action on vectors in \mathcal{H} . In particular we get selection rules and the Wigner–Eckart theorem, which are the subjects of the next section.

6.2 Selection Rules and the Wigner–Eckart Theorem

Now that we have defined representation operators, we can go on to formulate the fundamental selection rule from which the usual quantum mechanical selection rules can be derived. Then we can state and prove the Wigner–Eckart theorem, which is a kind of complement to the angular momentum selection rules. First, though, we need the following fact, the proof of which we only sketch. The details are deferred to the problems referenced below.

Proposition 6.1 *Let W_1 and W_2 be finite-dimensional inequivalent irreducible subspaces of a unitary representation (Π, \mathcal{H}) equipped with an inner product $(\cdot | \cdot)$. Then W_1 is orthogonal to W_2 .*

Proof sketch Define the **orthogonal projection operator** $P : \mathcal{H} \rightarrow W_2$ to be the map which sends $v \in \mathcal{H}$ to the *unique* vector $P(v) \in W_2$ satisfying

$$(P(v)|w) = (v|w) \quad \forall w \in W_2. \quad (6.8)$$

This is depicted schematically in Fig. 6.1. You will check in Problem 6.1 that such a vector $P(v)$ exists and is in fact unique. If we now restrict P to W_1 we get $P|_{W_1} : W_1 \rightarrow W_2$, and using the unitarity of Π and the invariance of the W_i one can show that $P|_{W_1}$ is an intertwiner. One can then use Schur’s lemma to conclude that $P|_{W_1} = 0$, which then implies that W_1 is orthogonal to W_2 . To fill out the details, see Problem 6.2. \square

With this in hand, we can now state and prove

Proposition 6.2 (Selection rule) *Let G be a semi-simple group, and let W_1 and W_2 be finite-dimensional inequivalent irreducible subspaces of a unitary representation (Π, \mathcal{H}) of G . Also let $\rho : V_0 \rightarrow \mathcal{L}(\mathcal{H})$ be a representation operator, where (Π_0, V_0) is some auxiliary representation. Then*

$$(w_1 | \rho(v) w_2) = 0 \quad \forall v \in V_0, w_i \in W_i \quad (6.9)$$

unless the decomposition of $V_0 \otimes W_2$ into irreducibles contains a representation equivalent to W_1 .

Before proving this, let us comment on what this somewhat opaque proposition really means. Roughly speaking, the selection rule just says that a vector of the form $\rho(v)w_2$ is a kind of ‘product’ of elements of V_0 and W_2 , and thus transforms like something in $V_0 \otimes W_2 = U_1 \oplus \cdots \oplus U_k$. Thus, for there to be any overlap with an irreducible subspace W_1 , W_1 must then be equivalent to one of the U_i .

Proof of proposition Define a map by

$$\begin{aligned} T : V_0 \otimes W_2 &\rightarrow \mathcal{H} \\ v \otimes w_2 &\mapsto \rho(v)w_2 \end{aligned} \quad (6.10)$$

and extending linearly to arbitrary elements of $V_0 \otimes W_2$. You can check that this is a linear map between vector spaces, and so the image $T(V_0 \otimes W_2) \equiv D \subset \mathcal{H}$ is a vector subspace of \mathcal{H} . Now, since G is semi-simple we can decompose $V_0 \otimes W_2$ into irreducibles as

$$V_0 \otimes W_2 = U_1 \oplus \cdots \oplus U_k$$

for some irreps U_k . As you will show below, the fact that ρ is a representation operator implies that T is in fact an intertwiner, and this further implies that the kernel of T (cf. Exercise 4.17) is an invariant subspace of $V_0 \otimes W_2$. This means that (with a possible relabeling of the U_i) we can write the kernel of T as $U_1 \oplus \cdots \oplus U_m$ for some $m \leq k$, which then implies that D is equivalent to $U_{m+1} \oplus \cdots \oplus U_k$. If none of the U_i are equivalent to W_1 , then by Proposition 6.1, every vector in D is orthogonal to every vector in W_1 , i.e.

$$(w_1 | \rho(v) w_2) = 0 \quad \forall v \in V_0, w_i \in W_i,$$

which is what we wanted to prove. \square

Exercise 6.1 Quickly show that ρ being a representation operator implies that T is an intertwiner. Show further that the kernel of T is an invariant subspace of $V_0 \otimes W_2$.

We will now use this generalized selection rule to reproduce some of the familiar selection rules from quantum mechanics.

Example 6.2 *Parity selection rules*

Let (Π, \mathcal{H}) be a complex unitary representation of the two-element group $\mathbb{Z}_2 = \{I, P\}$ where P is the parity operator. Let $v_\alpha, v_\beta \in \mathcal{H}$ be parity eigenstates with eigenvalues c_α, c_β . If $c_\alpha = 1$ then v_α spans the one-dimensional trivial representation of \mathbb{Z}_2 , and if $c_\alpha = -1$ then v_α spans the alternating representation. Likewise for c_β . Now let V_0 be another one-dimensional irrep of \mathbb{Z}_2 with parity eigenvalue c_0 , and let $B \in \rho(V_0)$, where ρ is a representation operator. Then a little thought shows that the selection rule implies

$$(v_\beta | B v_\alpha) = 0 \quad \text{unless } c_\beta = c_0 c_\alpha.$$

Thus if B is parity-odd ($c_0 = -1$) then it can only connect states of opposite parity, and if it is parity-even ($c_0 = +1$) then it can only connect states of the same parity. If we were looking at dipolar radiative transitions (which emit a photon) between electronic states of an atom or molecule, the relevant operators are the components of the dipole operator \mathbf{p} which is parity-odd (since it is a vector operator). The above then tells us that dipolar radiative transitions can only occur between electronic states of opposite parity.

Example 6.3 *Angular momentum selection rules*

Let (Π, \mathcal{H}) be a complex unitary representation of $SU(2)$, with two subspaces W_l and W_j equivalent to V_l and V_j , respectively, $2l, 2j \in \mathbb{Z}$. Also suppose we have an $SU(2)$ representation operator $\mathbf{A} : V_q \rightarrow \mathcal{L}(\mathcal{H})$, $q \in \mathbb{Z}$. Then for any $v \in W_l$, $v' \in W_j$, $A \in \mathbf{A}(V_q)$, the selection rule tells us that

$$(v' | A v) = 0 \quad \text{unless } |l - q| \leq j \leq l + q.$$

If we again consider a dipolar radiative transition between electronic states of an atom or molecule, the relevant operator is still the dipole \mathbf{p} whose components p_i live in $\rho(V_1)$, and so we find that a dipolar radiative transition between states with angular momentum j and l can only occur if $l - 1 \leq j \leq l + 1$. \square

The famous Wigner–Eckart theorem can be seen as a kind of complement to the angular momentum selection rule above. In the notation of the previous example, the Wigner–Eckart theorem says (roughly) that when $(v', A v)$ is *not* equal to zero, it is still tightly constrained and is in fact determined up to a constant by the fact that \mathbf{A} is a representation operator. The precise statement is as follows:

Proposition 6.3 (Wigner–Eckart) *Let (Π, \mathcal{H}) be a complex unitary representation of $SU(2)$, with two subspaces W_l and W_j equivalent to V_l and V_j , respectively, $2l, 2j \in \mathbb{Z}$. Also suppose we have two $SU(2)$ representation operators $\mathbf{A}, \mathbf{B} : V_q \rightarrow \mathcal{L}(\mathcal{H})$, $q \in \mathbb{Z}$, which yield two spherical tensors with components $A_k \equiv \mathbf{A}(v_k)$, $B_k \equiv \mathbf{B}(v_k)$, $0 \leq k \leq 2q + 1$. Finally, assume that*

$$(v' | A_k v) \neq 0 \quad \text{for some } k \text{ and } v \in W_l, v' \in W_j. \quad (6.11)$$

Then for all k and $w \in W_l, w' \in W_j$, we have

$$(w'|B_k w) = c(w'|A_k w) \quad (6.12)$$

for some constant $c \in \mathbb{C}$ which is independent of k, w and w' .

Proof Let $\mathcal{L}(W_l, W_j)$ denote the vector space of all linear maps from W_l to W_j . By restricting to W_l and using the orthogonal projection operator $P_j : \mathcal{H} \rightarrow W_j$, we can turn $\mathbf{A} : V_q \rightarrow \mathcal{L}(\mathcal{H})$ into a map

$$\begin{aligned} \tilde{\mathbf{A}} : V_q &\rightarrow \mathcal{L}(W_l, W_j) \\ v &\mapsto P_j \circ A(v)|_{W_l}. \end{aligned}$$

All we have done here is take the linear operator $A(v) \in \mathcal{L}(\mathcal{H})$ and restrict it to W_l and then project onto W_j . Now, since W_l, W_j are $SU(2)$ -invariant subspaces, the vector space $\mathcal{L}(W_l, W_j)$ actually furnishes a representation $(\tilde{\Pi}_1^1, \mathcal{L}(W_l, W_j))$ of $SU(2)$ by

$$\tilde{\Pi}_1^1(g)T \equiv \Pi(g)T\Pi(g)^{-1}, \quad T \in \mathcal{L}(W_l, W_j), \quad g \in SU(2). \quad (6.13)$$

You will check below that this is a bona fide representation, and is in fact equivalent to the tensor product representation $W_l^* \otimes W_j \simeq W_l \otimes W_j$! Furthermore, since the action of $SU(2)$ on \mathcal{H} commutes with restriction and projection (cf. Problem 6.2), it is not hard to see that $\tilde{\mathbf{A}}$ is an intertwiner. From (6.11) we know that $\tilde{\mathbf{A}}$ is not zero, and so from Schur's lemma we conclude that $\mathcal{L}(W_l, W_j) \simeq W_l \otimes W_j$ has a subspace U_q equivalent to V_q , and that $\tilde{\mathbf{A}}$ is a vector space isomorphism from V_q to U_q .

Now, we could also use our second representation operator \mathbf{B} to construct a second intertwiner $\tilde{\mathbf{B}} : V_q \rightarrow U_q \subset \mathcal{L}(W_l, W_j)$. Then we may invoke the corollary of Schur's lemma that you proved in Exercise 5.33 to conclude that $\tilde{\mathbf{B}} = c\tilde{\mathbf{A}}$. But this then means that

$$\begin{aligned} \tilde{\mathbf{A}}(v_k) &= c\tilde{\mathbf{B}}(v_k) \quad \forall k \\ \Rightarrow (w'|\tilde{\mathbf{A}}(v_k)w) &= c(w'|\tilde{\mathbf{B}}(v_k)w) \quad \forall k \text{ and } w \in W_l, w' \in W_j \\ \Rightarrow (w'|A_k w) &= c(w'|B_k w) \quad \forall k \text{ and } w \in W_l, w' \in W_j \end{aligned}$$

and so we are done. In deducing the last line we used the definition of $\tilde{\mathbf{A}}$, the definition of the orthogonal projection operator P_j , and the definitions $A_k = A(v_k)$, $B_k = B(v_k)$. \square

Exercise 6.2 Show that if $T \in \mathcal{L}(W_l, W_j)$ then so is $\Pi(g)T\Pi(g)^{-1}$, so that $(\tilde{\Pi}_1^1, \mathcal{L}(W_l, W_j))$ really is a representation; in fact, it is an invariant subspace of $(\Pi_1^1, \mathcal{L}(\mathcal{H}))$. In analogy to the equivalence between $V^* \otimes V$ and $\mathcal{L}(V)$, show that $(\tilde{\Pi}_1^1, \mathcal{L}(W_l, W_j))$ is equivalent to $W_l^* \otimes W_j$, which by Problem 5.3 is equivalent to $W_l \otimes W_j$.

You may have noticed that this is not the way the Wigner–Eckart theorem is usually stated in advanced quantum mechanics texts like Sakurai [14]. There, the theorem usually states that the matrix elements of a spherical tensor are proportional to the Clebsch–Gordan coefficients. To make this connection, consider the intertwiner

$$T : V_q \otimes W_l \rightarrow W_j$$

$$v \otimes w \mapsto P_j(\mathbf{A}(v)w).$$

If we work with standard bases $\{v_m\}_{m=0-2q}$, $\{w_n\}_{n=0-2l}$, $\{w'_p\}_{p=0-2j}$ for V_q , W_l , and W_j , then this map has components

$$\begin{aligned} T_{mn}^p &= T(v_m, w_n, w'^p) \\ &= w'^p(P_j(\mathbf{A}(v_m)w_n)) \\ &= \frac{1}{\|w'_p\|} (w'_p | P_j(A_m w_n)) \\ &= \frac{1}{\|w'_p\|} (w'_p | A_m w_n). \end{aligned} \tag{6.14}$$

Now let us switch to the orthonormal bases familiar from quantum mechanics, which look like

$$\begin{aligned} |q, m_q\rangle &\in V_q & -q \leq m_q \leq q \\ |l, m_l\rangle &\in W_l & -l \leq m_l \leq l \\ |j, m_j\rangle' &\in W_j & -j \leq m_j \leq j \end{aligned}$$

(notice the prime on the last set of vectors, which will distinguish it from vectors in $V_q \otimes W_l$ with the same quantum numbers). With this basis and notation, the components (6.14) of T become the matrix elements $\langle j, m_j | A_{m_q} | l, m_l \rangle$.

What do these matrix elements have to do with Clebsch–Gordan coefficients? Recall that $V_q \otimes W_l$ has two convenient sets of orthonormal basis vectors:

$$\begin{aligned} |q, m_q\rangle \otimes |l, m_l\rangle &\equiv |ql; m_q, m_l\rangle \in V_q \otimes W_l, & -q \leq m_q \leq q, & -l \leq m_l \leq l \\ |l', m_{l'}\rangle &\in V_q \otimes W_l, & |l - q| \leq l' \leq l + q, & -l' \leq m_{l'} \leq l'. \end{aligned}$$

The Clebsch–Gordan coefficients are just the inner products $\langle l', m_{l'} | ql; k, m_l \rangle$ of these basis vectors. Since $T : V_q \otimes W_l \rightarrow W_j$ is nonzero, $V_q \otimes W_l$ must contain a subspace U_j equivalent to W_j , and so we can consider the orthogonal projection operator $P : V_q \otimes W_l \rightarrow U_j$. By Problem 6.1 this is given by

$$P : V_q \otimes W_l \rightarrow U_j$$

$$|ql; m_q, m_l\rangle \mapsto \sum_{-j \leq m_j \leq j} \langle j, m_j | ql; m_q, m_l \rangle |j, m_j\rangle$$

and is an intertwiner by Problem 6.2. By then making the obvious identification of U_j with W_j and hence $|j, m_j\rangle \mapsto |j, m_j\rangle'$, we get the intertwiner

$$P' : V_q \otimes W_l \rightarrow W_j$$

$$|ql; m_q, m_l\rangle \mapsto \sum_{-j \leq m_j \leq j} \langle j, m_j | ql; m_q, m_l \rangle |j, m_j\rangle'$$

whose components are nothing but the Clebsch–Gordan coefficients! By Wigner–Eckart, though, this intertwiner must be proportional to T , and so its components

must be proportional to those of T . We thus have the following component version of the Wigner–Eckart theorem:

Proposition 6.4 (Wigner–Eckart II) *Let (Π, \mathcal{H}) be a complex unitary representation of $SU(2)$, with two subspaces W_l and W_j equivalent to V_l and V_j , respectively, $2l, 2j \in \mathbb{Z}$. Also suppose we have a degree q spherical tensor $\mathbf{A} = \{A_{m_q}\}$. Then*

$$\langle j, m_j | A_{m_q} | l, m_l \rangle = c \langle j, m_j | q l; m_q, m_l \rangle \quad (6.15)$$

where c is a constant independent of m_l, m_q , and m_j .

6.3 Gamma Matrices and Dirac Bilinears

We conclude this short chapter with what is essentially an extended example, which involves both the representation operators we have met in this chapter as well as the $O(3, 1)$ representation theory we developed in the last. This section relies on a familiarity with the theory of the Dirac electron; if you have not seen this material, and in particular are unfamiliar with gamma matrices and Dirac bilinears, then this section can be skipped.

Let $(D \equiv \Pi \oplus \bar{\Pi}, \mathbb{C}^4)$ be the Dirac spinor representation of $SL(2, \mathbb{C})$, so that

$$D : SL(2, \mathbb{C}) \rightarrow GL(4, \mathbb{C})$$

$$A \mapsto D(A) \equiv \begin{pmatrix} A & 0 \\ 0 & A^{\dagger-1} \end{pmatrix}.$$

We claim that the gamma matrices γ_μ can be seen as the components of a representation operator $\gamma : \mathbb{R}^4 \rightarrow \mathcal{L}(\mathbb{C}^4) = M_4(\mathbb{C})$, where \mathbb{R}^4 is the Minkowski four-vector representation of $SL(2, \mathbb{C})$. To define γ , we need the following two identifications of \mathbb{R}^4 with $H_2(\mathbb{C})$, where $X = (\mathbf{x}, t) \in \mathbb{R}^4$ and $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$:

$$\mathbb{R}^4 \leftrightarrow H_2(\mathbb{C})$$

$$X \leftrightarrow X_* \equiv \mathbf{x} \cdot \boldsymbol{\sigma} + tI \quad (6.16)$$

$$X \leftrightarrow X^* \equiv \mathbf{x} \cdot \boldsymbol{\sigma} - tI. \quad (6.17)$$

Note that (6.16) is just the identification we used in Examples 4.20 and 5.5 in defining the four-vector representation of $SL(2, \mathbb{C})$, and (6.17) is just a slight variation of that. Now, using the well-known property of the sigma matrices that

$$\sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij}, \quad (6.18)$$

you will verify below that

$$X^* = \eta(X, X) X_*^{-1}. \quad (6.19)$$

Also, if $\rho : SL(2, \mathbb{C}) \rightarrow SO(3, 1)_o$ is the homomorphism from Example 4.20 which defines the four-vector representation, then by the definition of ρ we have

$$(\rho(A)X)_* = AX_*A^\dagger, \quad (6.20)$$

which when combined with (6.19) yields

$$(\rho(A)X)^* = A^{\dagger-1}X^*A^{-1} \quad (6.21)$$

as you will also show below. If we then define a map γ by

$$\begin{aligned} \gamma : \mathbb{R}^4 &\rightarrow M_4(\mathbb{C}) \\ X &\mapsto \begin{pmatrix} 0 & X_* \\ X^* & 0 \end{pmatrix}, \end{aligned}$$

you can then use (6.20) and (6.21) to check that γ is a representation operator, i.e. that

$$\gamma(\rho(A)X) = D(A)\gamma(X)D(A)^{-1}. \quad (6.22)$$

If we define the gamma matrices as $\gamma_\mu \equiv \gamma(e_\mu)$, then we find that

$$\begin{aligned} \gamma_i &= \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad i = 1, 2, 3 \\ \gamma_4 &= \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \\ \gamma_5 &\equiv \gamma_1\gamma_2\gamma_3\gamma_4 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}. \end{aligned}$$

Up to a few minus signs that have to do with our choice of signature for the Minkowski metric, as well as the fact that most texts write γ_0 instead of γ_4 , this is the familiar *chiral* representation of the gamma matrices (the chiral representation is the one in which γ_5 is diagonal). If we let $\Lambda_\mu{}^\nu$ be the components of $\rho(A)$, we can take $X = e_\mu$ in (6.22) to get

$$\Lambda_\mu{}^\nu \gamma_\nu = D(A)\gamma_\mu D(A)^{-1} \quad (6.23)$$

which is how (6.22) usually appears in physics texts. Another important property is that the gamma matrices satisfy the fundamental anticommutation relation

$$\boxed{\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\eta_{\mu\nu}} \quad (6.24)$$

which you can easily verify.

Exercise 6.3 Verify (6.19), (6.21), (6.22), and (6.24).

Now we would like to construct the Dirac bilinears. Denote an element of \mathbb{C}^4 by

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

where ψ_L and ψ_R are two-component spinors living in $(\frac{1}{2}, \mathbf{0})$ and $(\mathbf{0}, \frac{1}{2})$, respectively. Also define the row vector $\bar{\psi}$ by

$$\bar{\psi} \equiv (\psi_R^*, \psi_L^*)$$

where the ‘ $*$ ’ denotes complex conjugation (note that the positions of the right and left-handed spinors are switched here). You can easily check that if ψ transforms like a Dirac spinor, i.e. $\psi \mapsto D(A)\psi$, then $\bar{\psi}$ transforms like

$$\bar{\psi} \mapsto \bar{\psi} \begin{pmatrix} A^{-1} & 0 \\ 0 & A^\dagger \end{pmatrix} = \bar{\psi} D(A)^{-1}. \quad (6.25)$$

If we consider the associated column vector $\bar{\psi}^T$, it transforms like

$$\bar{\psi}^T \mapsto \begin{pmatrix} A^{-1T} & 0 \\ 0 & A^* \end{pmatrix} \bar{\psi}^T$$

which you should recognize as the representation dual to the Dirac spinor, since the matrices of dual representations are just the inverse transpose of the original matrices. However, we know from Problem 5.2 that $(\frac{1}{2}, \mathbf{0})$ and $(\mathbf{0}, \frac{1}{2})$ are equivalent to their duals, so we conclude that $\bar{\psi}$ transforms like a Dirac spinor as well.

With $\bar{\psi}$ in hand, we can now define the *Dirac bilinears*

- $\bar{\psi}\psi$ scalar
- $\bar{\psi}\gamma_\mu\psi$ vector
- $\bar{\psi}\gamma_\mu\gamma_\nu\psi$, $\mu \neq \nu$ antisymmetric 2-tensor
- $\bar{\psi}\gamma_\mu\gamma_\nu\gamma_\rho\psi$, $\mu \neq \nu \neq \rho$ pseudovector
- $\bar{\psi}\gamma_5\psi$ pseudoscalar

Each of these contains a product of two Dirac spinors, and can be seen as components of tensors living in $\mathbb{C}^4 \otimes \mathbb{C}^4$. Note the similarities between the names of the Dirac bilinears and the first five entries of the table from Example 5.41. This makes it seem like the Dirac bilinears should transform like the components of antisymmetric tensors (of ranks 0 through 4). Is this true? Well, using (6.25), we find that the scalar transforms like

$$\bar{\psi} \mapsto \bar{\psi} D(A)^{-1} D(A) \psi = \bar{\psi} \psi$$

and so really does transform like a scalar. Similarly, using (6.23), we find that the vector transforms like

$$\begin{aligned} \bar{\psi}\gamma_\mu\psi &\mapsto \bar{\psi} D(A)^{-1} \gamma_\mu D(A) \psi \\ &= \Lambda^{-1}{}_\mu{}^\nu \bar{\psi}\gamma_\nu\psi \end{aligned}$$

and so really does transform like a vector. Using the anticommutation relation (6.24), you can similarly verify that the antisymmetric 2-tensor, pseudovector and pseudoscalar transform like antisymmetric tensors of ranks 2, 3, and 4, respectively. If we let $\Lambda^*(\mathbb{R}^4)$ denote the set of *all* antisymmetric tensor products of \mathbb{R}^4 , i.e.

$$\Lambda^*\mathbb{R}^4 \equiv \bigoplus_{k=0}^4 \Lambda^k\mathbb{R}^4,$$

and let $\Lambda^*\mathbb{R}_\mathbb{C}^4$ denote its complexification, then $\mathbb{C}^4 \otimes \mathbb{C}^4$ contains a subspace equivalent to $\Lambda^*\mathbb{R}_\mathbb{C}^4$. However, one can check that both spaces have (complex) dimension 16, so as $\mathfrak{sl}(2, \mathbb{C})_\mathbb{R}$ representations

$$\boxed{\mathbb{C}^4 \otimes \mathbb{C}^4 \simeq \Lambda^* \mathbb{R}_{\mathbb{C}}^4} \quad (6.26)$$

i.e. **the tensor product of the Dirac spinor representation with itself is equivalent to the space of all antisymmetric tensors!**

Another way to obtain this same result is to use our tensor product decomposition (5.95). The tensor product $\mathbb{C}^4 \otimes \mathbb{C}^4$ of two Dirac spinors is given by

$$\begin{aligned} & \left[\left(\frac{1}{2}, 0 \right) \oplus \left(0, \frac{1}{2} \right) \right] \otimes \left[\left(\frac{1}{2}, 0 \right) \oplus \left(0, \frac{1}{2} \right) \right] \\ &= \left[\left(\frac{1}{2}, 0 \right) \otimes \left(\frac{1}{2}, 0 \right) \right] \oplus \left[\left(\frac{1}{2}, 0 \right) \otimes \left(0, \frac{1}{2} \right) \right] \\ & \quad \oplus \left[\left(0, \frac{1}{2} \right) \otimes \left(\frac{1}{2}, 0 \right) \right] \oplus \left[\left(0, \frac{1}{2} \right) \otimes \left(0, \frac{1}{2} \right) \right] \\ &= (1, 0) \oplus (0, 0) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus (0, 1) \oplus (0, 0) \\ &= (0, 0) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus (1, 0) \oplus (0, 1) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus (0, 0) \end{aligned} \quad (6.27)$$

where in the first equality we used the fact that the tensor product distributes over direct sums (see Problem 5.11), in the second equality we used (5.95), and in the third equality we just rearranged the summands. However, from Example 5.41 we know that as an $\mathfrak{so}(3, 1)$ representation, $\Lambda^* \mathbb{R}_{\mathbb{C}}^4$ decomposes as

$$\boxed{\Lambda^* \mathbb{R}^4 \simeq (0, 0) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus (1, 0) \oplus (0, 1) \oplus \left(\frac{1}{2}, \frac{1}{2} \right) \oplus (0, 0)}$$

which is just (6.27)!

Exercise 6.4 Verify (6.25). Also verify that the antisymmetric 2-tensor, pseudovector and pseudoscalar bilinears transform like the components of antisymmetric tensors of rank 2, 3 and 4.

6.4 Problems

6.1 In this problem we will establish a couple of the basic properties of the orthogonal projection operator. To this end, let \mathcal{H} be a Hilbert space with inner product $(\cdot | \cdot)$ and let W be a finite-dimensional subspace of \mathcal{H} .

(a) Show that for any $v \in \mathcal{H}$, there exists a *unique* vector $P(v) \in W$ such that

$$(P(v)|w') = (v|w') \quad \forall w' \in W.$$

This defines the **orthogonal projection** map $P : \mathcal{H} \rightarrow W$ which projects \mathcal{H} onto W . (Hint: there are a few ways to show that $P(v)$ exists and is unique. One route is to consider the map $L : \mathcal{H} \rightarrow \mathcal{H}^*$ given by $L(v) = (v|\cdot)$ and then play around with restrictions to W .)

(b) Quickly show that $P(w) = w$ for all $w \in W$, and hence that $P^2 = P$.

- (c) Let $\{e_i\}$ be a (possibly infinite) orthonormal basis for \mathcal{H} where the first k vectors e_i , $i = 1, \dots, k$ are a basis for W . Show that if we expand an arbitrary $v \in \mathcal{H}$ as $v = v^i e_i$ where the implied sum is over *all* i (and where the sum may be infinite), then P takes the simple form

$$P(v) = P(v^i e_i) = \sum_{i=1}^k v^i e_i.$$

Thus P can be thought of as ‘projecting out all the components orthogonal to W ’.

- (d) Let $\{e_{j'}\}$ be an *arbitrary* ON basis for \mathcal{H} . Show that in this case the action of P is given by

$$P(e_{j'}) = \sum_{i=1}^k (e_i | e_{j'}) e_i.$$

6.2 Let (Π, \mathcal{H}) be a unitary representation of a group G on a Hilbert space \mathcal{H} . In this problem we will show that inequivalent irreducible subspaces of \mathcal{H} are orthogonal.

- (a) Let $W \subset \mathcal{H}$ be a finite-dimensional irreducible subspace and $P : \mathcal{H} \rightarrow W$ the orthogonal projection operator onto W . Use the defining property of P , the unitarity of Π , and the invariance of W to show that P is an intertwiner.
- (b) Let $V \subset \mathcal{H}$ be another irreducible subspace inequivalent to W . Restrict P to V to get $P|_V : V \rightarrow W$, which is still an intertwiner. Use Schur’s lemma to deduce that $P|_V = 0$, and conclude that V and W are orthogonal.

Appendix

Complexifications of Real Lie Algebras and the Tensor Product Decomposition of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ Representations

The goal of this appendix is to prove Proposition 5.8 about the tensor product decomposition of two $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representations. The proof is long but will introduce some useful notions, like the direct sum and complexification of a Lie algebra. We will use these notions to show that the representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ are in 1–1 correspondence with certain representations of the *complex* Lie algebra $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$, and the fact that this complex Lie algebra is a direct sum will imply certain properties about its representations, which in turn will allow us to prove Proposition 5.8.

A.1 Direct Sums and Complexifications of Lie Algebras

In this text we have dealt only with *real* Lie algebras, as that is the case of greatest interest for physicists. From a more mathematical point of view, however, it actually simplifies matters to focus on the complex case, and we will need that approach to prove Proposition 5.8. With that in mind, we make the following definition (in total analogy to the real case): A **complex Lie algebra** is a complex vector space \mathfrak{g} equipped with a complex-linear Lie bracket $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ which satisfies the usual axioms of antisymmetry

$$[X, Y] = -[Y, X] \quad \forall X, Y \in \mathfrak{g},$$

and the Jacobi identity

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \forall X, Y, Z \in \mathfrak{g}.$$

Examples of complex Lie algebras are $M_n(\mathbb{C}) = \mathfrak{gl}(n, \mathbb{C})$, the set of all complex $n \times n$ matrices, and $\mathfrak{sl}(n, \mathbb{C})$, the set of all complex, traceless $n \times n$ matrices. In both cases the bracket is just given by the commutator of matrices.

For our application we will be interested in turning real Lie algebras into complex Lie algebras. We already know how to complexify vector spaces, so to turn a real Lie algebra into a complex one we just have to extend the Lie bracket to the complexified vector space. This is done in the obvious way: given a real Lie algebra \mathfrak{g} with bracket

$[\cdot, \cdot]$, we define its **complexification** to be the complexified vector space $\mathfrak{g}_{\mathbb{C}} = \mathbb{C} \otimes \mathfrak{g}$ with Lie bracket $[\cdot, \cdot]_{\mathbb{C}}$ defined by

$$\begin{aligned} & [1 \otimes X_1 + i \otimes X_2, 1 \otimes Y_1 + i \otimes Y_2]_{\mathbb{C}} \\ & \equiv 1 \otimes [X_1, Y_1] - 1 \otimes [X_2, Y_2] + i \otimes [X_1, Y_2] + i \otimes [X_2, Y_1] \end{aligned}$$

where $X_i, Y_i \in \mathfrak{g}$, $i = 1, 2$. If we abbreviate $i \otimes X$ as iX and $1 \otimes X$ as X , this tidies up and becomes

$$[X_1 + iX_2, Y_1 + iY_2]_{\mathbb{C}} \equiv [X_1, Y_1] - [X_2, Y_2] + i([X_1, Y_2] + [X_2, Y_1]).$$

This formula defines $[\cdot, \cdot]_{\mathbb{C}}$ in terms of $[\cdot, \cdot]$, and is also exactly what you would get by naively using complex-linearity to expand the left hand side.

What does this process yield in familiar cases? For $\mathfrak{su}(2)$ we define a map

$$\begin{aligned} \phi : \mathfrak{su}(2)_{\mathbb{C}} &\rightarrow \mathfrak{sl}(2, \mathbb{C}) \\ 1 \otimes X_1 + i \otimes X_2 &\mapsto X_1 + \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} X_2, \quad X_1, X_2 \in \mathfrak{su}(2). \end{aligned} \tag{A.1}$$

You will show below that this is a Lie algebra isomorphism, and hence $\mathfrak{su}(2)$ complexifies to become $\mathfrak{sl}(2, \mathbb{C})$. You will also use similar maps to show that $\mathfrak{u}(n)_{\mathbb{C}} \simeq \mathfrak{gl}(n, \mathbb{R})_{\mathbb{C}} \simeq \mathfrak{gl}(n, \mathbb{C})$. If we complexify the real Lie algebra $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, we also get something nice. The complexified Lie algebra $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ has complex basis

$$\begin{aligned} \mathcal{M}_i &\equiv \frac{1}{2}(1 \otimes S_i - i \otimes \tilde{K}_i), \quad i = 1, 2, 3 \\ \mathcal{N}_i &\equiv \frac{1}{2}(1 \otimes S_i + i \otimes \tilde{K}_i), \quad i = 1, 2, 3 \end{aligned} \tag{A.2}$$

which we can again abbreviate¹ as

$$\begin{aligned} \mathcal{M}_i &\equiv \frac{1}{2}(S_i - i\tilde{K}_i), \quad i = 1, 2, 3 \\ \mathcal{N}_i &\equiv \frac{1}{2}(S_i + i\tilde{K}_i), \quad i = 1, 2, 3. \end{aligned}$$

These expressions are very similar to ones found in our discussion of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representations, and using the bracket on $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ one can verify that the analogs of (5.83) hold, i.e. that

$$\begin{aligned} [\mathcal{M}_i, \mathcal{N}_j]_{\mathbb{C}} &= 0 \\ [\mathcal{M}_i, \mathcal{M}_j]_{\mathbb{C}} &= \sum_{k=1}^3 \epsilon_{ijk} \mathcal{M}_k \\ [\mathcal{N}_i, \mathcal{N}_j]_{\mathbb{C}} &= \sum_{k=1}^3 \epsilon_{ijk} \mathcal{N}_k. \end{aligned} \tag{A.3}$$

¹Careful here! When we write $i\tilde{K}_i$, this is *not* to be interpreted as i times the matrix \tilde{K}_i , as this would make the \mathcal{N}_i identically zero (check!); it is merely shorthand for $i \otimes \tilde{K}_i$.

Notice that both $\text{Span}\{\mathcal{M}_i\}$ and $\text{Span}\{\mathcal{N}_i\}$ (over the complex numbers) are Lie subalgebras of $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ and that both are isomorphic to $\mathfrak{sl}(2, \mathbb{C})$, since

$$\text{Span}\{\mathcal{M}_i\} \simeq \text{Span}\{\mathcal{N}_i\} \simeq \mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C}).$$

Furthermore, the bracket between an element of $\text{Span}\{\mathcal{M}_i\}$ and an element of $\text{Span}\{\mathcal{N}_i\}$ is 0, by (A.3). Also, as a (complex) vector space $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ is the direct sum $\text{Span}\{\mathcal{M}_i\} \oplus \text{Span}\{\mathcal{N}_i\}$. When a Lie algebra \mathfrak{g} can be written as a direct sum of subspaces W_1 and W_2 , where the W_i are each subalgebras and $[w_1, w_2] = 0$ for all $w_1 \in W_1, w_2 \in W_2$, we say that the original Lie algebra \mathfrak{g} is a **Lie algebra direct sum** of W_1 and W_2 , and we write $\mathfrak{g} = W_1 \oplus W_2$.² Thus, we have the Lie algebra direct sum decomposition

$$\boxed{(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})}. \quad (\text{A.4})$$

This decomposition will be crucial in our proof of Proposition 5.8.

Exercise A.1 Prove that (A.1) is a Lie algebra isomorphism. Remember, this consists of showing that ϕ is a vector space isomorphism, and then showing that ϕ preserves brackets. Then find similar Lie algebra isomorphisms to prove that

$$\begin{aligned} \mathfrak{u}(n)_{\mathbb{C}} &\simeq \mathfrak{gl}(n, \mathbb{C}) \\ \mathfrak{gl}(n, \mathbb{R})_{\mathbb{C}} &\simeq \mathfrak{gl}(n, \mathbb{C}). \end{aligned}$$

A.2 Representations of Complexified Lie Algebras and the Tensor Product Decomposition of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ Representations

In order for (A.4) to be of any use, we must know how the representations of a real Lie algebra relate to the representations of its complexification. First off, we should clarify that when we speak of a representation of a complex Lie algebra \mathfrak{g} we are ignoring the complex vector space structure of \mathfrak{g} ; in particular, the Lie algebra homomorphism $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is only required to be **real-linear**, in the sense that $\pi(cX) = c\pi(X)$ for all *real* numbers c . If \mathfrak{g} is a complex Lie algebra, the representation space V is complex and $\pi(cX) = c\pi(X)$ for all *complex* numbers c , then we say that π is **complex-linear**. Not all complex representations of complex Lie algebras are complex-linear. For instance, all of the $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ representations described in Sect. 5.10 can be thought of as representations of the complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$, but only those of the form $(\mathbf{j}, \mathbf{0})$ are complex-linear, as you will show below.

²Notice that this notation is ambiguous, since it could mean either that \mathfrak{g} is the direct sum of W_1 and W_2 as vector spaces (which would then tell you nothing about how the direct sum decomposition interacts with the Lie bracket), or it could mean Lie algebra direct sum. We will be explicit if there is any possibility of confusion.

Exercise A.2 Consider all the representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ as representations of $\mathfrak{sl}(2, \mathbb{C})$ as well. Show directly that the fundamental representation $(\frac{1}{2}, \mathbf{0})$ of $\mathfrak{sl}(2, \mathbb{C})$ is complex-linear, and use this to prove that all $\mathfrak{sl}(2, \mathbb{C})$ representations of the form $(\mathbf{j}, \mathbf{0})$ are complex-linear. Furthermore, by considering the operators N_i defined in (5.82), show that these are the *only* complex-linear representations of $\mathfrak{sl}(2, \mathbb{C})$.

Now let \mathfrak{g} be a *real* Lie algebra and let (π, V) be a complex representation of \mathfrak{g} . Then we can extend (π, V) to a complex-linear representation of the complexification $\mathfrak{g}_{\mathbb{C}}$ in the obvious way, by setting

$$\pi(X_1 + iX_2) \equiv \pi(X_1) + i\pi(X_2), \quad X_1, X_2 \in \mathfrak{g}. \quad (\text{A.5})$$

(Notice that this representation is complex-linear by definition, and that the operator $i\pi(X_2)$ is only well-defined because V is a *complex* vector space.) Furthermore, this extension operation is reversible: that is, given a complex-linear representation (π, V) of $\mathfrak{g}_{\mathbb{C}}$, we can get a representation of \mathfrak{g} by simply restricting π to the subspace $\{1 \otimes X + i \otimes 0 \mid X \in \mathfrak{g}\} \simeq \mathfrak{g}$, and this restriction reverses the extension just defined. Furthermore, you will show below that (π, V) is an irrep of $\mathfrak{g}_{\mathbb{C}}$ if and only if it corresponds to an irrep of \mathfrak{g} . We thus have

Proposition A.1 *The irreducible complex representations of a real Lie algebra \mathfrak{g} are in one-to-one correspondence with the irreducible complex-linear representations of its complexification $\mathfrak{g}_{\mathbb{C}}$.*

This means that we can identify the irreducible complex representations of \mathfrak{g} with the irreducible complex-linear representations of $\mathfrak{g}_{\mathbb{C}}$, and we will freely make this identification from now on. Note the contrast between what we are doing here and what we did in Sect. 5.9; there, we complexified real representations to get complex representations which we could then classify; here, the representation space is fixed (and is always complex!) and we are complexifying the *Lie algebra itself*, to get a representation of a complex Lie algebra on the same representation space we started with.

Exercise A.3 Let (π, V) be a complex representation of a real Lie algebra \mathfrak{g} , and extend it to a complex-linear representation of $\mathfrak{g}_{\mathbb{C}}$. Show that (π, V) is irreducible as a representation of \mathfrak{g} if and only if it is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$.

Example A.1 *The complex-linear irreducible representations of $\mathfrak{sl}(2, \mathbb{C})$*

As a first application of Proposition A.1, consider the complex Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. Since $\mathfrak{sl}(2, \mathbb{C}) \simeq \mathfrak{su}(2)_{\mathbb{C}}$, we conclude that its complex-linear irreps are just the irreps (π_j, V_j) of $\mathfrak{su}(2)$! In fact, you can easily show directly that the complex-linear $\mathfrak{sl}(2, \mathbb{C})$ representation corresponding to (π_j, V_j) is just $(\mathbf{j}, \mathbf{0})$. \square

As a second application, note that by (A.4) the complex-linear irreps of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ are just the representations $(\mathbf{j}_1, \mathbf{j}_2)$ coming from $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$. Since $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ is a direct sum, however, there is another way to construct complex-linear irreps. Take two complex-linear irreps of $\mathfrak{sl}(2, \mathbb{C})$, say (π_{j_1}, V_{j_1})

and (π_{j_2}, V_{j_2}) . We can take a modified tensor product of these representations such that the resulting representation is not of $\mathfrak{sl}(2, \mathbb{C})$ but rather of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$. We denote this representation by $(\pi_{j_1} \bar{\otimes} \pi_{j_2}, V_{j_1} \otimes V_{j_2})$ and define it by

$$\begin{aligned} (\pi_{j_1} \bar{\otimes} \pi_{j_2})(X_1, X_2) &\equiv \pi_{j_1}(X_1) \otimes I + I \otimes \pi_{j_2}(X_2) \in \mathcal{L}(V_1 \otimes V_2), \\ X_1, X_2 &\in \mathfrak{sl}(2, \mathbb{C}). \end{aligned} \quad (\text{A.6})$$

Note that we have written the tensor product in $\pi_{j_1} \bar{\otimes} \pi_{j_2}$ as ‘ $\bar{\otimes}$ ’ rather than ‘ \otimes ’; this is to distinguish this tensor product of representations from the tensor product of representations defined in Sect. 5.3. In the earlier definition, we took a tensor product of two \mathfrak{g} representations and produced a third \mathfrak{g} representation given by $(\pi_1 \otimes \pi_2)(X) = \pi_1(X) \otimes I + I \otimes \pi_2(X)$, where the *same* element $X \in \mathfrak{g}$ gets fed into both π_1 and π_2 ; here, we take a tensor product of two \mathfrak{g} representations and produce a representation of $\mathfrak{g} \oplus \mathfrak{g}$, where two *different* elements $X_1, X_2 \in \mathfrak{g}$ get fed into π_1 and π_2 .

Now, one might wonder if the representation $(\pi_{j_1} \bar{\otimes} \pi_{j_2}, V_{j_1} \otimes V_{j_2})$ of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ defined above is equivalent to $(\mathbf{j}_1, \mathbf{j}_2)$; this is in fact the case! To prove this, recall the following notation: the representation space $V_{j_1} \otimes V_{j_2}$ is spanned by vectors of the form $v_{k_1} \otimes v_{k_2}$, $k_i = 0, \dots, 2j_i$, $i = 1, 2$ where the v_{k_i} are characterized by (5.76). Similarly, the representation space $V_{(j_1, j_2)}$ of $(\mathbf{j}_1, \mathbf{j}_2)$ is spanned by vectors of the form v_{k_1, k_2} , $k_i = 0, \dots, 2j_i$, $i = 1, 2$, where these vectors are characterized by (5.87). We can thus define the obvious intertwiner

$$\begin{aligned} \phi : V_{j_1} \otimes V_{j_2} &\rightarrow V_{(j_1, j_2)} \\ v_{k_1} \otimes v_{k_2} &\mapsto v_{k_1, k_2}. \end{aligned}$$

Of course, we must check that this map actually is an intertwiner, i.e. that

$$\phi \circ [(\pi_{j_1} \bar{\otimes} \pi_{j_2})(X_1, X_2)] = \pi_{(j_1, j_2)}(X_1, X_2) \circ \phi \quad \forall X_1, X_2 \in \mathfrak{sl}(2, \mathbb{C}). \quad (\text{A.7})$$

Since $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ is of (complex) dimension six, it suffices to check this for the six basis vectors

$$\begin{aligned} (i\mathcal{M}_z, 0) \\ (0, i\mathcal{N}_z) \\ (\mathcal{M}_+, 0) &\equiv (i\mathcal{M}_x - \mathcal{M}_y, 0) \\ (\mathcal{M}_-, 0) &\equiv (i\mathcal{M}_x + \mathcal{M}_y, 0) \\ (0, \mathcal{N}_+) &\equiv (0, i\mathcal{N}_x - \mathcal{N}_y) \\ (0, \mathcal{N}_-) &\equiv (0, i\mathcal{N}_x + \mathcal{N}_y), \end{aligned}$$

where the \mathcal{M}_i and \mathcal{N}_i were defined in (A.2). We now check (A.7) for $(i\mathcal{M}_z, 0)$ on an arbitrary basis vector $v_{k_1} \otimes v_{k_2}$, and leave the verification for the other five $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ basis vectors to you. The left hand of (A.7) gives (careful with all the parentheses!)

$$\begin{aligned}
\phi((\pi_{j_1} \otimes \pi_{j_2})(i\mathcal{M}_z, 0)(v_{k_1} \otimes v_{k_2})) &= \phi((J_z \otimes I)(v_{k_1} \otimes v_{k_2})) \\
&= \phi((j_1 - k_1)v_{k_1} \otimes v_{k_2}) \\
&= (j_1 - k_1)v_{k_1, k_2}
\end{aligned}$$

while the right hand side is

$$\begin{aligned}
\pi_{(j_1, j_2)}(i\mathcal{M}_z, 0)(\phi(v_{k_1} \otimes v_{k_2})) &= (\pi_{(j_1, j_2)}(i\mathcal{M}_z, 0))(v_{k_1, k_2}) \\
&= iM_z v_{k_1, k_2} \\
&= (j_1 - k_1)v_{k_1, k_2}
\end{aligned}$$

and so they agree. The verification for the other five $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ basis vectors proceeds similarly. This proves the following proposition:

Proposition A.2 *Let $(\mathbf{j}_1, \mathbf{j}_2)$ denote both the usual $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irrep as well as its extension to a complex-linear irrep of $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$. Let $(\pi_{j_1} \bar{\otimes} \pi_{j_2}, V_{j_1} \otimes V_{j_2})$ be the representation of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ defined in (A.6). Then*

$$(\mathbf{j}_1, \mathbf{j}_2) \simeq (\pi_{j_1} \bar{\otimes} \pi_{j_2}, V_{j_1} \otimes V_{j_2}). \quad (\text{A.8})$$

Exercise A.4 Verify (A.7) for the other five $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$ basis vectors, and thus complete the proof of Proposition A.2.

We will now use Proposition A.2 to compute tensor products of the $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ irreps $(\mathbf{j}_1, \mathbf{j}_2)$, which by Proposition A.1 will give us the tensor products of the $(\mathbf{j}_1, \mathbf{j}_2)$ as $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps, which was what we wanted! In the following computation we will use the fact that

$$(\pi_{j_1} \bar{\otimes} \pi_{j_2}) \otimes (\pi_{k_1} \bar{\otimes} \pi_{k_2}) \simeq (\pi_{j_1} \otimes \pi_{k_1}) \bar{\otimes} (\pi_{j_2} \otimes \pi_{k_2}), \quad (\text{A.9})$$

which you will prove below (note which tensor product symbols are ‘barred’ and which are not). With this in hand, and using the $\mathfrak{su}(2)_{\mathbb{C}} \simeq \mathfrak{sl}(2, \mathbb{C})$ tensor product decomposition (5.68), we have (omitting the representation spaces in the computation)

$$\begin{aligned}
(\mathbf{j}_1, \mathbf{j}_2) \otimes (\mathbf{k}_1, \mathbf{k}_2) &\simeq (\pi_{j_1} \bar{\otimes} \pi_{j_2}) \otimes (\pi_{k_1} \bar{\otimes} \pi_{k_2}) \\
&\simeq (\pi_{j_1} \otimes \pi_{k_1}) \bar{\otimes} (\pi_{j_2} \otimes \pi_{k_2}) \\
&\simeq \left(\bigoplus_{l_1=|j_1-k_1|}^{j_1+k_1} \pi_{l_1} \right) \bar{\otimes} \left(\bigoplus_{l_2=|j_2-k_2|}^{j_2+k_2} \pi_{l_2} \right) \\
&\simeq \bigoplus_{(l_1, l_2)} \pi_{l_1} \bar{\otimes} \pi_{l_2} \quad \text{where } |j_i - k_i| \leq l_i \leq j_i + k_i, \quad i = 1, 2 \\
&\simeq \bigoplus_{(l_1, l_2)} (\mathbf{l}_1, \mathbf{l}_2) \quad \text{where } |j_i - k_i| \leq l_i \leq j_i + k_i, \quad i = 1, 2.
\end{aligned}$$

This gives the tensor product decomposition of complex-linear $(\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}})_{\mathbb{C}}$ irreps. However, these irreps are just the extensions of the irreps $(\mathbf{j}_1, \mathbf{j}_2)$ of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, and

you can check that the process of extending an irrep to the complexification of a Lie algebra commutes with taking tensor products, so that the extension of a product is the product of an extension. From this we conclude the following:

Proposition A.3 *The decomposition into irreps of the tensor product of two $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps $(\mathbf{j}_1, \mathbf{j}_2)$ and $(\mathbf{k}_1, \mathbf{k}_2)$ is given by*

$$(\mathbf{j}_1, \mathbf{j}_2) \otimes (\mathbf{k}_1, \mathbf{k}_2) = \bigoplus (l_1, l_2)$$

$$\text{where } |j_1 - k_1| \leq l_1 \leq j_1 + k_1, \quad |j_2 - k_2| \leq l_2 \leq j_2 + k_2$$

which is just Proposition 5.8.

Exercise A.5 Prove (A.9) by referring to the definitions of both kinds of tensor product representations and by evaluating both sides of the equation on an arbitrary vector $(X_1, X_2) \in \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$.

References

1. V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer, Berlin, 1989
2. L. Ballentine, *Quantum Mechanics: A Modern Development*, World Scientific, Singapore, 1998
3. A. Cannas Da Silva, *Lectures on Symplectic Geometry*, Lecture Notes in Mathematics 1764, Springer, Berlin, 2001
4. T. Frankel, *The Geometry of Physics*, 1st ed., Cambridge University Press, Cambridge, 1997
5. S. Gasiorowicz, *Quantum Physics*, 2nd ed., Wiley, New York, 1996
6. H. Goldstein, *Classical Mechanics*, 2nd ed., Addison-Wesley, Reading, 1980
7. M. Gökeler and T. Schücker, *Differential Geometry, Gauge Theories, and Gravity*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1987
8. B. Hall, *Lie Groups, Lie Algebras and Representations: An Elementary Introduction*, Springer, Berlin, 2003
9. I. Herstein, *Topics in Algebra*, 2nd ed., Wiley, New York, 1975
10. K. Hoffman and D. Kunze, *Linear Algebra*, 2nd ed., Prentice Hall, New York, 1971
11. A.L. Onishchik, *Lectures on Real Semisimple Lie Algebras and Their Representations*, ESI Lectures in Mathematics and Physics, 2004
12. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, San Diego, 1972
13. W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York, 1976
14. J.J. Sakurai, *Modern Quantum Mechanics*, 2nd ed., Addison-Wesley, Reading, 1994
15. B. Schutz, *Geometrical Methods of Mathematical Physics*, Cambridge University Press, Cambridge, 1980
16. S. Sternberg, *Group Theory and Physics*, Princeton University Press, Princeton, 1994
17. V.S. Varadarajan, *Lie Groups, Lie Algebras and Their Representations*, Springer, Berlin, 1984
18. F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Springer, Berlin, 1979

Index

A

Abelian (group), 88
 Active transformation, 50
 ad homomorphism, 132, 136
 Ad homomorphism, 136
 Addition of angular momentum, 64, 65
 Additive quantum number, 160, 161
 Adjoint, 92
 Adjoint representation, 21, *see* representation, adjoint
 of Lorentz group, 149, 175
 of orthogonal group, 149, 175
 Alternating tensor, 70
 Angular momentum, 6, 128
 commutation relations, 122
 Angular momentum operators, *see* operators, angular momentum
 Angular velocity vector, 6, 78, 123
 Anti-Hermitian matrices, 118
 Anti-Hermitian operators
 Lie algebra of, 127
 Antisymmetric
 2nd rank tensor representation, 149
 matrices, 16, 116, 118
 tensor, 70
 tensors and Lie algebras of isometry groups, 177
 Axial vector, 75
 Axis of rotation, 74, 138

B

Baker–Campbell–Hausdorff formula, 117
 Basis (for a vector space), 15
 Bilinear, 27
 Bivector, 75
 Body axes, 19
 Body frame, 78

Boost, 100, 125, 139
 generators, 119

Bosons, 71, 111

Bra, 32

C

$\mathcal{C}(P)$, 128, 158
 \mathbb{C}^n , 11
 as complexification of \mathbb{R}^n , 194
 Canonical
 transformation, 129
 variables, 131
 Cartesian product, xv
 Chiral representation (of gamma matrices), 223
 Clebsch–Gordan coefficients, 66, 221
 Clifford algebra, 208
 Closed (group), 112
 Cofactor expansion of determinant, 73
 Commutator, 116
 Completely reducible, 180
 Completeness (of vector space), 29
 Complex Lie algebra, 227
 Complex vector space, 11
 Complex-linear, 229
 Complexification
 of Lie algebra, 227
 of representation, 194
 of vector space, 194
 Components
 contravariant, 32, 46
 covariant, 32, 46
 of dual vector, 25
 of matrix, 22
 of Minkowski metric, 29
 of tensor, 4, 40, 55
 of vector, 17

Conjugate representation, *see* representation, conjugate
 Conjugate-linear, 27
 Connected group, 98
 Contraction, 56
 Contravariant, *see* components, contravariant
 Covariant, *see* components, covariant
 Cross product, 76
 Cyclic permutation, 72

D

Decomposable, 180
 Determinant, 73, 82, 83
 and oriented volume, 74, 83
 and sgn homomorphism, 110
 of a matrix exponential, 119
 Dimension (of a vector space), 15
 Dirac bilinear, 224
 Dirac delta functional, 33, 61
 Dirac notation, 32, 40, 57, 61, 66
 Dirac spinor, 147, 151, 183, 205, 222, 224
 and antisymmetric tensors, 225
 Direct sum, 178, 183
 external, 183
 internal, 183
 of Lie algebras, 229
 representation, 183
 Direct sum decomposition
 of $\Lambda^2 \mathbb{R}^4$, 181
 of $L^2(S^2)$, 182
 of $M_n(\mathbb{C})$, 179
 of $M_n(\mathbb{R})$, 179
 of tensor product representations, 180
 Disconnected group, 98
 Dot product, 28
 Double cover
 of $SO(3)$ by $SU(2)$, 107
 of $SO(3, 1)_0$ by $SL(2, \mathbb{C})$, 108
 Double dual, 35
 Dual
 basis, 26
 metric, 36
 representation, 161, 162
 space, 25
 vector, *see* linear functional

E

ϵ , *see* Levi-Civita tensor
 Einstein summation convention, 39, 46
 Electromagnetic field tensor, 60
 transformation properties of, 149, 181, 203
 Entanglement, 67
 Equivalence of representations, 169
 Euler angles, 97

Euler's theorem, 138
 Exponential of matrix, 114

F

Faithful representation, 210
 Fermions, 71, 111
 Four-vector, 147
 Four-vector representation, *see* representation, four-vector
 Fourier coefficients, 20
 Fourier series, 20
 Fourier transform, 62
 Fundamental representation, *see* representation, fundamental

G

$GL(n, \mathbb{R})$, *see* general linear group
 $GL(V)$, *see* general linear group
 $\mathfrak{gl}(n, \mathbb{C})$, 115
 $\mathfrak{gl}(V)$, 127
 Gamma matrices, 222, 223
 chiral representation of, 223
 General linear group, 14, 90, 115
 Generators, 111
 Gram–Schmidt, 27
 Group, 81
 axioms, 88
 definition of, 88

H

$H_n(\mathbb{C})$, *see* Hermitian matrices
 $\mathcal{H}_l(\mathbb{R}^3)$, *see* harmonic polynomials
 $\tilde{\mathcal{H}}_l$, *see* spherical harmonics
 Hamilton's equations, 158
 Harmonic polynomials, 12, 16, 23, 34, 48, 69, 155, 192
 Heisenberg algebra, 130, 157, 160
 Heisenberg equation of motion, 21
 Heisenberg picture, 52
 Hermite polynomials, 37
 Hermitian adjoint, 35
 Hermitian matrices, 11, 16
 Hermitian scalar product, 28
 Hermiticity, 27
 Highest weight vector, 189, 197
 Hilbert space, 17, 29
 Hilbert–Schmidt norm (on $M_n(\mathbb{C})$), 117
 Homomorphism
 group, 103
 Lie algebra, 131

I

Identical particles, 71, 150

Indices

raising and lowering, 32

Inner product, 27

on \mathbb{C}^n , 28

on $L^2([-a, a])$, 29

on $M_n(\mathbb{C})$, 28

on \mathbb{R}^n , 28

Inner product space, 27

Intertwiner, 169

Intertwining map, *see* intertwiner

Invariant subspace, 179

non-trivial, 179

Inversion, 75, 98

Invertibility, 22, 83

Irreducible representation, 180, 184, 185

of a real Lie algebra and its

complexification, 230

of abelian group, 187

of double-cover of $O(3, 1)$, 205

of $O(3, 1)$, 206–208

of $SL(2, \mathbb{C})$, 201

of $SO(3)$, 192

of $SO(3, 1)_o$, 201

of $\mathfrak{sl}(2, \mathbb{C})$, 230

of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, 196–201

of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$, 230–232

of $\mathfrak{su}(2)$, 188–190

of \mathbb{Z}_2 , 187

Irrep, *see* irreducible representation

Isometry, 90, 104

Isomorphism

group, 103

Lie algebra, 131

J

J_+ , 188, 191, 192

J_- , 188, 191, 192, 194

J_z , 188, 191, 192, 194

Jacobi identity, 116, 133

K

Kernel, 105

Ket, 32

Killing form, 137, 143

Kinetic energy of rigid body, 6

L

$L^2([-a, a])$, 12, 16, 20, 28, 32, 60

$L^2(\mathbb{R})$, 12, 60

$L^2(\mathbb{R}^3)$, 64, 160

$L^2(S^2)$, 155, 182

$\mathcal{L}(V)$, *see* operators, linear

Laguerre polynomials, 37

Left-handed spinor, *see* representation, spinor

Legendre polynomials, 37

Levi-Civita symbol, *see* Levi-Civita tensor

Levi-Civita tensor, 41, 72, 110, 208, 212

Lie, Sophus, 111

Lie algebra, 111, 115–131

abstract, 127

complex, 227

definition, 115

direct sum, 229

physicist's definition, 119

Lie algebra homomorphism, 131

induced, 133

Lie algebra representation, 145

Lie bracket

and similarity transformation, 137

Lie group, 111–113

dimension of, 112

Lie product formula, 116, 134

Lie subalgebra, 128

Linear combinations, 14

Linear dependence, independence, 15

Linear functional, 25

Linear map, 104

Linear operator

as $(1, 1)$ tensor, 39, 57

definition of, 21

Lie algebra of, 127

Linearity, 21

Lorentz force law, 60

Lorentz group, 93, 99–102

Lorentz transformations, 49, 99

decomposition into boost and rotation, 139

M

$M_n(\mathbb{C})$, 227

as complexification of $M_n(\mathbb{R})$, 194

$M_n(\mathbb{R})$, $M_n(\mathbb{C})$, 11, 16

Matrix

exponential, 114

logarithm, 142

of linear operator, 23

of Minkowski metric, 29

Matrix Lie group, 112

Maxwell stress tensor, 59

Metric, 27

as intertwiner between V and V^* , 173

Minkowski metric, 28, 93

Moment of inertia tensor, 6, 42, 58

Momentum representation, 62

Multilinearity, 3, 39

Multiplicative quantum number, 160, 161

Multipole moments, 42, 69

N

Neutrino, 147
 Noether's theorem, 129
 Non-abelian (group), 88
 Non-commutative (group), 88
 Non-degeneracy, 27
 Non-degenerate Hermitian form, 27
 Null space, 106
 Nullity, 106

O

$O(3)$, 98, 115
 $O(3, 1)$, 101, 138
 $o(3)$, 115
 $o(n - 1, 1)$, 119
 $o(n)$, 118
 $O(n - 1, 1)$, *see* Lorentz group
 $O(n)$, *see* orthogonal matrices, orthogonal group
 Observables, 128
 One-parameter subgroup, 114
 One-to-one, 22
 Onto, 22
 Operators
 adjoint of, 35
 angular momentum, 21, 23
 exponential of, 34
 Hermitian, 36
 invertible, 22
 linear, *see* linear operator
 matrix representation of, 23
 self-adjoint, 36
 symmetric, 36
 Orientation, 74, 83
 Orthogonal
 complement, 210
 matrices, 47, 81
 projection operator, 217, 225
 set, 28
 Orthogonal group, 91, 96–98
 proof that it is a matrix Lie group, 112
 Orthonormal basis, 27
 for Hilbert space, 30

P

$P_l(\mathbb{C}^2)$, 152, 185, 191
 $P_l(\mathbb{R}^3)$, *see* polynomials, of degree l
 \hat{p} , 21, 60, 62, 130
 Parity, 101, 160, 188
 and Dirac spinors, 183
 and \mathbb{Z}_2 , 109
 Passive transformation, 50, 81
 Pauli matrices, 12

Permutation

 even, odd, 70, 110
 Permutation group, 95, 150
 relation to \mathbb{Z}_2 , 109
 Permutations, *see* permutation group
 Phase space, 128
 Pin groups, 208
 Poisson bracket, 128
 formulation of mechanics, 158
 Polar decomposition theorem, 140
 Polynomials, 12
 harmonic, *see* harmonic polynomials
 Hermite, 37
 Laguerre, 37
 Legendre, 37
 real, 36
 Positive matrix, 140
 Positive-definite, 27
 Principal minors, 140
 Product
 of matrices, 23
 Product state, 67
 Pseudoscalar, 167, 168
 Pseudovector, 75, 149, 168, 182

Q

\hat{q} , 130
 Quark, antiquark, 173

R

Rank (of map), 106
 Rank (of tensor), 3, 40
 Rank-nullity theorem, 106
 Rapidity, 100
 Real numbers (as additive group), 89
 Real vector space, 11
 Real-linear, 229
 Representation, 128, 145
 2nd rank antisymmetric tensor, 149, 175, 203
 2nd rank antisymmetric tensor and adjoint of $O(n)$, 166
 adjoint, 148
 adjoint of $C(P)$, 158
 alternating, 150
 complex, 145
 conjugate, 209
 dual, 161, 162, 173
 equivalence of, 169
 faithful, 210
 four-vector, 147, 151, 201
 fundamental, 147
 irreducible, 180
 of Heisenberg algebra on $L^2(\mathbb{R})$, 157

Representation (*cont.*)

- of S_n , 150
- of \mathbb{Z}_2 , 150
- on function spaces, 151
- on linear operators, 163
- on symmetric and antisymmetric tensors, 165–169
- on $T_s^r(V)$, 161
- pseudoscalar, 168
- pseudovector, 168
- real, 145
- scalar, 166
- sgn, 150
- $SO(2)$ on \mathbb{R}^2 , 187, 196
- $SO(3)$ on $\mathcal{H}_l(\mathbb{R}^3)$, $\tilde{\mathcal{H}}_l$ and $L^2(S^2)$, 155
- $SO(3)$ on $L^2(\mathbb{R}^3)$, 153
- space, 145
- spin s , 152, 160
- spin-one, 195
- spin-two, 195
- spinor, 147, 151, 174, 198, 199, 205, 222
- $SU(2)$ on $P_l(\mathbb{C}^2)$, 152
- symmetric traceless tensors, 195
- tensor product, 159
- trivial, 146
- unitary, 146
- vector, 147
 - and “spin-one”, 194

Representation operator, 216

Right-handed spinor, *see* representation, spinor

Rigid body motion, 19, 52, 78

Rotation, 74, 96–98, 129, 138

- generators, 118, 119, 121, 122
- improper, 98
- proper, 98

S

$S^l(\mathbb{C}^2)$, 165, 192, 199, 200

S_n , *see* permutation group

$S^r(V)$, *see* symmetric, tensors

$SL(2, \mathbb{C})$, 102, 151

- relationship to $SO(3, 1)_o$, 108, 135

$SO(2)$, 96

$SO(3)$, 97, 106, 129

- infinitesimal elements of, 113

$SO(3, 1)_o$, 99, 138

$SO(n)$, *see* special orthogonal group

$SU(2)$, 98

- relationship to $SO(3)$, 106, 135, 138

$SU(3)$, 173

$SU(n)$, *see* special unitary group

$\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$, 125

- relationship to $\mathfrak{so}(3, 1)$, 136

$\mathfrak{sl}(n, \mathbb{C})$, 227

$\mathfrak{so}(2)$, 121

$\mathfrak{so}(3)$, 122, 129

$\mathfrak{so}(3, 1)$, 124

$\mathfrak{so}(n)$, 120

$\mathfrak{su}(2)$, 106, 124

- isomorphic to $\mathfrak{so}(3)$, 135

$\mathfrak{su}(n)$, 120

Scalar representation, *see* representation, scalar

Scalars, 9

Schrödinger picture, 52

Schur’s lemma, 186

Selection rule, 218

- parity, 219

Self-adjoint, 36

Semi-simple (group or algebra), 180

Separable state, 67

Sgn homomorphism, 110

Sgn representation, 150

Similarity transformation, 48, 164

Space axes, 19

Space frame, 78

Span (of a set of vectors), 14

Special linear group, 102

Special orthogonal group, 94, 138

Special unitary group, 94, 98

Spectral theorem, 156

Spherical harmonics, 12, 16, 23, 34, 155, 182

Spherical Laplacian, 13, 156

Spherical tensor, 214, 216

Spin, 11, 64

Spin angular momentum, 65

Spin-one representation, *see* representation, spin-one

Spin-two representation, *see* representation, spin-two

Spinor, *see* representation, spinor

Spinor representation, *see* representation, spinor

Square-integrable, 12, 156

Star operator, 212

Structure constants, 132

Subgroup, 90

Subspace, 11

Symmetric

- form, 27
- matrices, 16
- tensors, 68

Symmetric group, *see* permutation group

Symmetrization postulate, 111, 150

T $[T]_{\mathcal{B}}$, 23 \mathcal{T}_s^r , *see* tensors, of type (r, s)

Tensor operator, 214

Tensor product

as addition of degrees of freedom, 63–67, 160

of operators, 63

of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{R}}$ irreps, 202, 233of $\mathfrak{su}(2)$ irreps, 181

of vector spaces, 54

of vectors, 54

representation, 159

Tensors

alternating, 70

and linear operators, 5, 57

and matrices, 5

antisymmetric, 70

as linear operators, 39

as tensor product space, 55

basis for vector space of, 55

components of, 4, 40, 55

contraction of, 56

definition of, 39

rank of, 40

symmetric, 68

transformation law, 5, 43–53

type (r, s) , 39, 55

Time-reversal, 102

and \mathbb{Z}_2 , 109

Trace, 26, 48

cyclic property of, 121

interpretation of, 120

Transformation law

of linear operators, 48, 163

of metric tensors, 48

of vectors and dual vectors, 46, 162

Translations, 130

Transpose, 35

Transposition, 68, 109

Trivial representation, *see* representation, trivial**U** $U(1)$, 105 $U(n)$, *see* unitary group $u(n)$, 118

Unitary

group, 92

matrices, 47

operator, 92

representation, 146

V V^* , *see* dual space

Vector operators, 63, 213, 214

Vector representation, *see* representation, vector

Vector space

as additive group, 90

axioms, 9

complex, 11

definition of, 9

isomorphism, 105

real, 11

W

Wedge product, 70, 110

Weyl spinor, 147

Wigner–Eckart theorem, 219–222

X \hat{x} , 21, 60**Y** Y_m^l , *see* spherical harmonics**Z** \mathbb{Z}_2 , 95, 150

and parity, time-reversal, 109

and sgn homomorphism, 109

irreducible representations of, 187

is semi-simple, 187, 210